THE JOURNAL OF PARAPSYCHOLOGY

Volume 16 SEPTEMBER, 1952 Number 3

Editorial PAGE

What Kind of Interest Sustains the Psi Researches? 157

Some Relations Between Call Sequence and ESP Performance .. 165
J. G. Pratt and S. G. Soal

Why Throw Dice? 187
R. A. McConnell

Minor Articles

Effect of Novelty in Test Conditions on ESP Performance .. 192
Remi J. Cadoret

A Study of the Effect of Conditioning on Multiple-Aspect ESP Scoring 204
Stanley J. Hallett

Effect of the Receiver's Attitude Toward the Sender in ESP Tests 212
George W. Casper

Notes and Comments

Experiments in ESP and PK with Aboriginal Subjects 219
Ronald Rose

Glossary 221

A SCIENTIFIC QUARTERLY DEALING WITH EXTRASENSORY PERCEPTION, THE PSYCHOKINETIC EFFECT, AND RELATED TOPICS

DUKE UNIVERSITY PRESS
DURHAM, N. C.

The Journal *of* Parapsychology

EDITORIAL STAFF

J. B. RHINE and J. G. PRATT, *Editors*
ELIZABETH A. McMAHAN, *Assistant Editor*
J. A. GREENWOOD and T. N. E. GREVILLE, *Statistical Editors*
DOROTHY H. POPE, *Managing Editor*

ADVISORY EDITORIAL BOARD

HORNELL HART
Duke University

H. H. PRICE
University of Oxford

GARDNER MURPHY
College of the City of New York

ROBERT H. THOULESS
University of Cambridge

This JOURNAL is published on the fifteenth day of March, June, September, and December by Duke University Press, Durham, North Carolina.

Contributions submitted for publication and all editorial communications should be addressed to the Managing Editor, Dorothy H. Pope, College Station, Durham, North Carolina. Correspondence with the editors is advised before submitting articles other than reports of experimentation. Since it does not forward manuscripts by registered mail, the JOURNAL cannot guarantee that they will not be lost in transit, and contributors are urged to keep copies of their papers. All contributions should be typewritten, double spaced. References should be given in the form adopted by the JOURNAL.

Reprints may be ordered when the proof is returned. The JOURNAL will bear one half the cost of reprints up to two hundred copies. Correspondence concerning subscriptions, change of address, back numbers, and other business communications should be addressed to Dorothy H. Pope, College Station, Durham, N. C.

The subscription price is $6.00 a year; single current numbers $1.50. The rate for back volumes is $7.00; for single numbers $1.75. Missing numbers will be supplied free when lost in the mails if written notice is given within one month of the date of issue. All remittances should be made payable to the JOURNAL OF PARAPSYCHOLOGY.

Subscriptions may be obtained in Great Britain through the Cambridge University Press, Bentley House, Euston Road, London, N. W. 1.

Entered as second class matter May 21, 1942, at the post office at Durham, N. C., under the act of March 3, 1879.

DUKE UNIVERSITY PRESS

The Journal *of*
Parapsychology

Volume 16 SEPTEMBER, 1952 Number 3

EDITORIAL

WHAT KIND OF INTEREST SUSTAINS THE PSI RESEARCHES?

IN ASKING what interest it is that sustains the researches in parapsychology, we need to consider two different groups of people. First there is the all-too-small group of investigators that actually do the research. We may include in this group all of those who are concerned with the activities leading up to the actual work and with bringing the findings to the interested reader. Even in this inclusive conception of the group, there are not more than a few score of individuals who are at any given time actively engaged in what most of the readers of this JOURNAL would consider acceptable or useful research on psi problems.

In the history of every branch of science there comes a point when the research workers are numerous enough to constitute an adequate professional audience; but although parapsychology has made progress in recent decades, it has not reached the point where there are enough parapsychologists to make up an adequate supporting group. As a matter of fact, the number of active workers is so small that they could not hope to accomplish very much in so difficult a field on the strength of their interest and efforts alone. They must have the support of some clientele, some interested section of the general population.

The other, and fortunately much larger, group is the one with which this editorial is chiefly concerned. It is made up of all those

persons who actively follow the publication of findings with sufficient interest to do something in consequence. If they do nothing more than to subscribe to a periodical, buy a book, write a letter, or report a personal psi experience, they come in this class of responsive readers.

If we may judge from letters, subscriptions, book sales, and the like and from the impressions of editors, it seems safe to infer that this group numbers hundreds of thousands around the world, most of its members being concentrated in Western Europe and North America, but with considerable numbers in India, Japan, Australia, New Zealand, South Africa, and South America. This group, which we will call the psi audience, provides the encouragement, the recognition, and the various kinds of support (represented most concretely, though not exclusively, in terms of finances) that sustain the efforts of the active workers in the field.

What is the basic interest that activates this psi audience? What common aim sustains these individuals in their support of the psi researches? One of the elements of this interest no doubt is the simple, natural curiosity we have about strange things. A somewhat larger share of interest springs from the need of many persons to explain their own psychic experiences. There is, however, a more basic and lasting interest than these that distinguishes this group and leads its members to give serious consideration to the attempts being made to solve the problems of parapsychology by sound research. There would have to be some enduring type of interest to have sustained such researches through the years and even over the decades.

Let us look at the matter in its historical setting. The main interest that activated the founders of parapsychology was the challenge psi phenomena offered to the materialism of nineteenth-century science. This was particularly true of the intellectual leadership in Britain, where, at the time, interest in such phenomena was the most vigorous. Accordingly, there was a general readiness in the psychical research societies to turn to the claims of spiritualism and the phenomena of mediumship for the most convincing answer to the challenge of materialism. The prospect of establishing the existence of a world of incorporeal personalities seemed to afford the best

possible refutation of a physical theory of the nature of man. But from the initial high pitch of interest in spiritistic claims that existed during the last quarter of the century there has been a very marked decline. There are today many fewer people ready to investigate such phenomena and still fewer ready to contribute adequately to the help of such investigations. Some interest does still exist, but it appears to be hardly strong enough to sustain a proper research program.

The experiments in telepathy, however, which were a part of the work of the early societies at the same time that inquiries were being made into the phenomena of mediumship, have had a different sort of history. Gradually the quality and coverage of these experiments have developed, until now the research findings under the general heading of psi phenomena have reached a degree of soundness and a measure of recognition never approximated in the mediumistic studies. These investigations of psi capacities by strictly experimental methods have won a large measure of respect from our psi audience and even from a still wider public that is interested but somewhat less concerned over the problems than our sustaining group.

What, then, is it about these psi phenomena that brings the research this vitally important support? While it is necessarily somewhat conjectural to speak for such a clientele, we think we have grounds to justify the following impression of the central interest in these psi phenomena. It is that they assure us of the presence of something that expands rather than contracts the conception of human personality, something that extends rather than restricts the vision of man's place in nature, something that suggests wider potentialities for human life rather than reduces them, something, finally, that gives some support from science for the conception of man on which our social values and institutions have been founded.

It is this basic problem of the nature of man that forms the central core of interest identifying the psi audience. Probably for most of its members this interest is not very explicitly thought out. For some it may be a dim, half-unconscious impression, a mild excitement, a slender hope, a faint flutter of encouragement to offset the depressing conception of man inherent in the mechanistic dog-

mas. But whether formulated or not, it is, we think, this concern with man's basic nature that defines the character of this group and inspires its attention to the psi inquiries.

In a word, the interest of the psi audience is of the same kind, generally speaking, as that which has activated the workers themselves from the beginning of parapsychology as a scientific study.

What, then, can we say of the make-up of the psi audience? What types of people respond with the kind of interest we have indicated as the central one in parapsychology? Can they be classified under any familiar headings, or identified with any recognizable group?

We definitely need to answer this general question, if only for ourselves as workers in the field. In this precarious stage of pioneering, with its demands, difficulties, and uncertainties, we ought to know more about our friends and allies. And they in turn, spread all over the world as they are, may wish to know something about the group to which they belong. Otherwise each member may be far too unaware of the degree to which his own interests are shared and supported by others.

It would be supposed from the nature of the problems with which parapsychology is concerned that the psi audience would be made up mainly of psychologists and psychiatrists. Also, since psi phenomena offer a factual basis for a refutation of materialism, it might be expected that they would be of outstanding interest to the leadership of the organized religions. Or again, many persons would anticipate finding that the psi audience consists chiefly of the occult groups, the fringe of cults and creeds and faiths that have sprung up in the vacuum, the no-man's-land that lies between the orthodox religions and the conventional professions concerned with the same problem areas. Let us briefly examine these possibilities.

Unfortunately we cannot say that our audience is primarily psychological. There are now, it is true, many psychologists in this country and abroad who are taking the psi researches very seriously, some of them even doing active investigations in their own laboratories. Many others are introducing the problems and results to their students in the classroom and in the laboratory. Some are making quiet, personal investigations for their own satisfaction, and

a growing number of psychology students are finding out by extra-curricular reading something about the work in parapsychology. There is now comparatively little open criticism from psychologists, and no new critical issue has been raised in many years.

Nevertheless, we cannot say that any substantial part of the psi audience is made up of psychologists. The swing of psychologists toward an active interest in the psi researches is comparatively slow, and this slow rate of change is understandable. Parapsychology is not just another school to take its place along with current psychological theories and practices. Its acceptance by psychologists in general will probably require a more basic theoretical shift for their branch of science than any other innovation in its history. A departure from the ideological framework of the physical sciences to which psychology has clung for so many decades will require for most psychologists an inordinate amount of evidence and (what is more effective) more time for the slow processes of familiarization to take effect.

The story of the relation of parapsychology to the psychiatrists, however, is more encouraging. Psychiatry, less concerned with the ultimate categorization of its basic doctrine and more hard-pressed by the practical exigencies of life and mental illness, is a bolder branch of inquiry. Historically, psychology is indebted to the greater venturesomeness of the psychiatrist for many of its basic concepts. Spurred by his immediate contact with the real problems of mankind, he has made his laboratory a less shielded one. Proportionately, then, more of the psi audience is psychiatric than psychological. In the active participation of professional men in psi researches, psychiatrists are numerically in the lead, and there is every indication from such sources as correspondence, lectureships, surveys of professional reaction, and the like that a large percentage of psychiatrists read and listen to the progressive developments in parapsychology. Yet psychiatrists, too, are only a small minority in the sustaining group of responsive psi readers.

What of the various types of professional religionists? Is it from these that we draw our main support? The answer is emphatically negative. The psi researches receive no conspicuous support either from the various formal religions or from the occult

organizations. With the exception of a vigorous and promising movement that is now developing in one of the most liberal denominations, there has been almost no response or expression from any of the religious groups as such.

The fact is understandable. The leadership of these groups is selected to preserve the doctrines of the organization and not to invest in new explorations or extensions. The attitude encouraged is one of frank conservatism. A spirit of inquiry would be contrary to the manifest objectives and values of the group. Recognition of a need for the kind of support that parapsychology might offer to religious doctrine would be considered by many religious leaders as a confession of weakness, a lack of faith that would be intolerable. There has, however, been active support for parapsychology by many individuals from among religious professions, just as among scientific ones. From many a pulpit, efforts have been made to bring the underpinning value of the psi investigations to the support of the doctrines of religion, more especially to the reinforcement of religion's case against a physicalistic view of man.

It would, however, be utterly wrong to say that our present supporting group is a conspicuously religious one. It is religious only in the sense that some of the problems to which the religions offer answers are, broadly speaking, the central problems being investigated by parapsychologists: the question of the ultimate nature of man, the possibility of a spiritual factor in human life, the problem whether man has the power and freedom to lift and control himself. Naturally our psi audience is especially concerned about the scientific findings on these same problems that have already been given religious solutions and interpretations, but this audience is preponderantly a lay group in the strict sense of the word.

The fact is, our psi audience is not easily classified in any ordinary way, being primarily neither scientific nor religious, and not linked in any way with a particular system or organization. It simply is made up of those who are to some extent going along with us in our efforts to explore these great unknowns in the life of man. The group that sustains parapsychology today is not characterizable other than as the alert reading public that is reached through the daily press, magazines, books, radio, and television. Its

distribution cuts across all the professions and classes, races and creeds, and includes a wide range of educational backgrounds.

The realization of the character of this sustaining clientele brings us to a consideration of the ways in which the group has grown and the ways in which our contact with it is maintained. For the most part this acquaintanceship has been accomplished through that broad avenue of communication, the science press, with strong secondary support from the magazines of national and international coverage. Those persons responsible for interpreting science for the general public probably have done more to make it possible for the researches in parapsychology to continue and to develop than any other single factor to which we can turn. The fact that these avenues of education have been kept open to such new and revolutionary findings, that there has been a high degree of editorial and reportorial fairness, clarity, impartiality, and efficiency has in this instance been in the interest of science itself. For science as a whole needs a public, needs free access to it, and needs to bypass the censorship that is potentially inherent in editorial decisions. Parapsychology has encountered this editorial censorship many times, especially in professional scientific journals. It has even on occasion been accompanied by a frank admission of bias.

The benefits that come from this large supporting psi audience can only be touched on. There are case reports, contact with gifted subjects, financial contributions, research suggestions, and helpful ideas of various kinds. The greater coverage enables us to bring to the research a range of professional assistance that could not be purchased by a financially limited branch of inquiry. Public interest unlocks many doors and provides many opportunities. It is easier to get help for an undertaking in which hundreds of thousands are already interested, and it is harder for such an inquiry to be ignored or repressed with a world looking on than it would otherwise be. And finally, it is encouraging to the lone worker, who has little else perhaps besides his own interest to sustain him, to know that there is a fair-sized world of people who would also like to know what the answer to his research question is going to turn out to be.

We in the research field are bound to this audience with a sincere fraternal feeling, as common interests naturally bind people. Its

members, mostly unknown individually, form a large but still very real public audience that we know are our genuine sustaining friends. They are friends who will criticize as well as praise, who will appreciate not only our successes when they come but also the difficulties of attaining them, and friends, we believe, whose support will enable research in parapsychology to continue until its objectives have been attained. J. B. R.

SOME RELATIONS BETWEEN CALL SEQUENCE AND ESP PERFORMANCE

By J. G. Pratt and S. G. Soal[1]

ABSTRACT: This paper continues previous analyses made on the data of Mrs. Gloria Stewart, the subject in extensive ESP experiments carried out by Dr. S. G. Soal. The previous article dealt with the relation between success in hitting a specified target and the pattern of targets in which the hit occurred. The present paper deals with the success in hitting the target as related to the pattern of calls in which the successful call was situated.

In the first paper it was found that the subject scored below chance expectation on trials involving sequences of identical targets. The later analyses have shown that this was true only when the calls for those targets were not alike. To satisfy certain statistical requirements, the analyses described in this paper were based, as were most of those of the earlier study, upon non-overlapping pairs of trials. When the calls as well as the targets were doubled, hits on the first target were scored in 25.1 percent of all pairs. When the targets alone were doubled, hits were made on the first target in only 16.2 percent of all pairs. The difference between the two cases was statistically significant. The writers interpret this finding to mean that the subject, who was shown in the previous study to have some kind of tendency not to hit target doubles, did not always know by ESP when a particular target was a member of a double, and on those trials she scored positively, just as she did on target singles.

A second relationship between call sequence and scoring to be investigated was the frequency of repeating a symbol in order to score a hit and repeating a symbol when making a miss. A significantly larger proportion of call doubles were found in the former case, though this trend was not present in all sections of the data. The interpretation offered is that the subject, who had a general aversion to calling doubles, was able to overcome this aversion to some degree in scoring some of her ESP hits.

Finally, the writers investigated the frequency of repeating a call symbol on the trial following a hit as compared with the frequency of repeating a call symbol after a miss. It was found that there was a significantly higher number of call doubles in the former case, and no interpretation for this fact could be offered other than that the subject was somehow aware of or sensitive to the fact that she had made an ESP hit and that this fact was registered in her call sequence.—Ed.

I. Introduction

This is the second report we have made upon variations in ESP success in the work of Mrs. Gloria Stewart. The study of these data has revealed that there were factors operating from trial to trial within the ESP tests which produced highly significant differences

[1] J.G.P.'s participation in this joint research was made possible through the financial support of the Rockefeller Foundation, and that of S.G.S. through a Fulbright award.

in performance. We are concerned in our reports solely with these secondary ESP effects; that is, with effects that show up only as the result of special analyses and which cannot be detected simply by adding up the hits made on a succession of trials.

The ESP records with which these investigations are concerned are those from the experiments that S.G.S. and his collaborators conducted during the period 1945-50 with the subject, Mrs. Gloria Stewart. In these tests, the subject wrote her call on each trial for one of five animals that was designated as the target by a sender in the next room. Further details regarding the experimental conditions and general results were given in our previous report (**5**).

In our preceding paper we presented results which showed that Mrs. Stewart's ESP response was not simply a matter of reacting to one of five animal names that had been designated as the target for a given trial. Rather, the subject's success or failure in choosing the correct animal initial was sometimes influenced by the targets of the preceding and following trials as well as by the one designated for the immediate trial. The highest success was obtained when the target was unlike either the preceding one or the following one. In fact, it appeared from the point reached in the analyses of the earlier report that only in this instance was the subject able to make more than the chance number of hits, as she actually scored significantly below the chance level when the target at which she was aiming was preceded or followed by one or more trials with the same animal initial as target. In other words, below-chance scoring was associated with target doubles, triples, or longer unbroken sequences. The low scoring on sequences of identical targets was not simply an absence of ESP success; it was actually ESP in reverse to a significant degree.

Research workers in parapsychology have long been aware that there may be factors in the experimental test situation which cause the subject to score below the chance level when he is trying to score positively. Only now, however, has the extent of the psi-missing effect been recognized as fully as it is in the paper by Dr. Rhine in the last number of the JOURNAL (**2**). The example of ESP missing with which we have to deal in Mrs. Stewart's records is not exactly typical, but it may be that the very aspects in which her

negative scoring is atypical may throw further light on the factors concerned in psi-missing. Most of the examples of negative scoring to which Dr. Rhine calls attention have taken the form of a below-chance deviation in the total score of an individual percipient or a group of percipients working under certain experimental conditions. In Mrs. Stewart's data, however, we find the positive scoring and the negative scoring taking place in the closest possible proximity, with the subject shifting from ESP hitting of the target on one trial to ESP missing of the next target which follows after only two or three seconds.

We have felt, therefore, that it is important to make every reasonable effort to find out as much as we can about the peculiarities of scoring in this subject's work. Our primary object, of course, is to go as far as we can at the present time toward understanding the factors which influenced her ESP performance. It will then remain to be seen whether the factors at work were purely individual ones that apply only to Mrs. Stewart's performance or whether they can be generalized to the work of other subjects. Even if we cannot interpret her results fully at this time, the facts are clearly worth recording.

The results that we wish to discuss in this paper relate to three general topics, but before these are introduced it will be advantageous to review the basic analysis of our earlier report. For that study of ESP performance in relation to target sequence, we selected for analysis 1,340 columns or runs from Mrs. Stewart's records. These represented all of her work under the standard test conditions in which the sender knew the target, except for trials done after her scores declined to the chance level at the end of the five-year series. The details regarding experimental conditions and selection of data are fully covered in the preceding report and need not be reviewed here. Suffice it to say that in the basic analysis the data of each column were divided into twelve non-overlapping pairs of trials, the twenty-fifth trial in each column being ignored. This gave 16,080 pairs of trials for analysis, and these were further separated into pairs with both targets the same (doubles) and pairs with unlike targets (non-doubles). We then compared the rate of scoring on the first target of each pair within the two groups, target doubles

and target non-doubles, and found that there was a highly significant difference, with positive scoring on target singles and negative scoring on target doubles. The rest of the preceding report was devoted to extending and refining this general result, but the review that has been given is all that is required for an understanding of the present study.

At the time the analysis of the target sequence effect was carried out, we tabulated the information regarding each pair of trials by means of a scheme not mentioned in the preceding report. Two additional facts were tabulated with regard to each pair of trials besides those described above, and these provided the data of the present study. These facts were: (a) each pair of trials was classified according to whether the subject's calls were double or non-double; and (b) the pairs were also grouped according to whether the second trial gave a hit or a miss.

With the help of this fresh information we are able to answer in this report three further questions regarding Mrs. Stewart's ESP performance. (1) Was the subject's getting a negative deviation on the first trial of pairs with target doubles related in any way to the fact that the two calls of the pair were double or non-double? (2) Did the subject depart from her usual calling sequence, as shown by the proportion of doubles called, when she scored a hit on the second member of the target pair? (3) Was there any difference in the subject's tendency to repeat a call symbol after making a hit on the first trial of a pair and after making a miss on the first trial? As these questions deal with more or less distinct psychological problems, it will be more convenient to present and discuss the results bearing on them separately.

II. The Relation of Call Doubles to the ESP Scoring on Target Doubles and Non-Doubles

Results

The first step toward finding out whether the call sequence was related to the negative scoring effect found on target doubles was to separate all the pairs with target doubles into two groups, those with and those without call doubles. The two groups were then arranged in a four-fold table to show the number of hits on the first target and the number of misses on the first target of the pair in

each group as shown in Table 1. The upper figure in each cell shows the observed cases falling in the indicated category, while the lower figure (in parentheses) indicates chance expectation based on the given marginal totals of the table for the observed data. A chi-square test applied to this table revealed that there was a highly significant

Table 1

RATE OF SCORING ON THE FIRST MEMBERS OF NON-OVERLAPPING PAIRS OF TRIALS WITH TARGET DOUBLES

	Pairs with Call Doubles	Pairs with Call Non-Doubles	Total
Hits on First Trial	(*a*) 97 (66.5)	(*b*) 466 (496.5)	563
Misses on First Trial	(*c*) 289 (319.5)	(*d*) 2418 (2387.5)	2707
Total	386	2884	3270
Ratio of Hits to Total	25.1%	16.2%	17.2%

$$x^2 = \frac{(ad-bc)^2\,(a+b+c+d)}{(a+b)\,(c+d)\,(a+c)\,(b+d)} = 19.22 \text{ (1 d.f.)}$$

$$P = .000012$$

difference in the rate of scoring in the two cases. The chi square of 19.22 indicates that the difference is so large we could expect it to occur by chance only once in approximately 100,000 cases. The number of hits on the first member of target doubles in those pairs in which the subject called doubles is 25.1 percent of the total number of pairs with both target and call doubles, while only 16.2 percent of the pairs with target doubles alone show hits on the first trial.

To complete the study of the effect of call doubles, we made a similar comparison between hits on the first member of pairs with and without call doubles for those pairs in which the targets were non-double (Table 2). In this analysis no significant difference in rate of scoring was found between call doubles and non-doubles. In both instances the pairs with hits on the first trial were found to be 26.8 percent of the total pairs in the group.

This division of the data in relation to the call sequence has

Table 2

RATE OF SCORING ON THE FIRST MEMBERS OF NON-OVERLAPPING PAIRS OF TRIALS WITH TARGET NON-DOUBLES

	Pairs with Call Doubles	Pairs with Call Non-Doubles	Total
Hits on First Trial	(*a*) 354 (353.8)	(*b*) 3077 (3077.2)	3431
Misses on First Trial	(*c*) 967 (967.2)	(*d*) 8412 (8411.8)	9379
Total	1321	11489	12810
Ratio of Hits to Total	26.8%	26.8%	26.8%

$\chi^2 = .00015$ (1 d.f.) P = .99+

therefore revealed a very interesting fact. It has shown that the low scoring on the first member of target doubles is confined to those pairs in which the calls were not doubled. The probability that any particular call will hit the target for that trial is statistically independent of the next call. On a purely chance basis, therefore, the expected ratio of hits to total trials in all four columns of Tables 1 and 2 is one to five, or 20 percent. The 25.1 percent success found when a pair contains both call doubles and target doubles is not significantly different from the 26.8 percent found when the targets are non-doubles.

Discussion

What interpretation is it possible to make of the fact that the subject scored significantly below chance on the first member of a target double *only when the two calls of the pair were different from each other, or non-double*? We concluded from the general results summarized in our previous report that Mrs. Stewart for some reason used her ESP ability *as a general rule* to score negatively or to miss the target when she was presented with two or more identical targets in succession. For this rule to take effect, however, it was necessary for her to detect *by ESP* that the targets were identical. She had, in other words, to perceive them by ESP as a double, triple, quadruple, etc., but her ESP did not work infallibly

in this regard any more than it served without fail to identify the isolated targets (singles) on which she scored positively.

When the subject's ESP failed to let her know that the target she was responding to was a member of a double or longer sequence, that does not mean that her psi function had no opportunity to operate. She could still use her ESP on each separate target just as if it were a single or isolated one. Thus, her very failure to put the target configuration by ESP in the class of those she ordinarily responded to by psi-missing left her free to respond positively to each target as if it were an isolated one.

Now if the subject failed to appreciate the fact that two like targets actually formed a double and that therefore she ought to avoid hitting them, the coming into play of her natural tendency to hit those targets would tend to result in her calling doubles. By isolating those cases in which doubles in the calls are associated with doubles in the targets we have brought together in one group those instances in which the subject failed to recognize by ESP that the targets were double, and this gives us a special class of target doubles in which her ordinary positive scoring tendencies prevailed.

If we are correct in this interpretation, we have an example of a particular situation or pattern in the target sequence which generally produced negative scoring, but which did not always do so. We appear to be dealing here with one of those rare instances in which we are able to demonstrate statistically what investigators have suggested many times on the basis of clinical impressions. This is that there is much more evidence for ESP in our test results than is brought out by ordinary methods. Usually we have to be satisfied with a net result in which positive scoring and negative scoring trends may partly cancel each other out. Apparently in this case we have successfully separated the opposing trends and have found a special condition (call doubles) in which the score on target doubles was *above* the chance level although for this subject it was usually below.

III. Frequency of Calling Doubles or Non-Doubles When Scoring a Hit or a Miss

Problem

The question with which we are concerned in this section has to

do with the extent to which the subject must break away from his habitual call sequences in order to make an ESP hit. We know that if a person is asked to make up lists using a prescribed set of objects, such as many columns or runs of ESP symbols or of five letters representing animal names, these subjectively formed sequences will not be random distributions. This will be true whether a person making up such a subjective sequence is consciously attempting to imitate randomness or is simply calling off the symbols without any effort to control or direct his associative processes. In either case, he will soon fall into sequence habits which will be his personal pattern of calling and which will tend to persist over a period of time, though they may change with experience.

Sequence habits in calling are also to be found in any ESP test in which the subject is not hitting at virtually the 100 percent scoring level, and for practical purposes this includes all ESP experiments. The presence of these non-random call patterns implies a tendency for each call to be determined by preceding responses instead of by the ESP target for a particular trial. If the subject is to succeed in making a call that corresponds with the target instead of one that simply follows by habit from the preceding calls, it is logical to suppose that he must break away from his normal associative tendencies. We would expect to find, therefore, in studying any particular feature of the call sequence, such as the number of doubles, that there would be a difference in frequency of occurrence between those calls that score hits on the target and those that make misses. In fact, we might go further and predict that the number of doubles completed by calls that score hits should come closer to the expected chance frequency of doubling in a random series than do those calls that make misses. A successful call series is necessarily determined to some extent by the target series and so would tend to reflect certain of its random features.

The foregoing prediction assumes that the break in the subject's sequence habits occurs precisely at the point of scoring an ESP hit. This assumption that the subject goes on following his usual associative tendencies until he gets an ESP impression strong enough to enable him to break through his sequence habits is perhaps the simplest one. We must recognize, however, that this is not the only

way in which impressions can be superimposed upon the subject's own peculiar subjective pattern of calling. It is conceivable that the ESP target is effective in influencing the subject's choice of calls one, two, or even more trials before the one on which the actual ESP hit is scored. The subject might, in other words, know in advance that a particular target symbol is coming up, say, on the third trial ahead and this ESP knowledge might influence him to choose at that point one of his habitual sequence patterns which will end up on the third call of the sequence with the symbol that will score a hit. If ESP should operate in this anticipatory fashion by influencing the subject's choice among his habitual call patterns, the selections might be made most easily at points in his associative chain of responses where the links are weak. We should then have the ESP process operating so as to take advantage of the subject's habitual sequences, and not having to work against them by operating only at the moment the target is presented.

The question, then, is whether the call sequences are broken to a statistically significant degree precisely at the point of making a hit. We would expect this to be so, but this is no reason for neglecting an opportunity to test the point. Up to the present, ESP has produced so many surprises that we should not hesitate to investigate a question merely because the answer appears to be obvious.

In confining ourselves to the frequency of *doubles* in the calls, we are of course restricting our attention to only one of a large number of call sequences that might be examined. The present study, therefore, is merely an introduction to the general problem of the relation of ESP performance to the subject's personal associative tendencies. A more exhaustive investigation, though tedious and time-consuming, might throw new light on the mechanism by which ESP impressions are brought to the level of conscious response.

Results

In this study of doubles and non-doubles in the calls in relation to whether the second member of the call pair scored a hit or a miss, it was necessary to classify the 16,080 non-overlapping pairs of trials in two ways. First, did the second trial of the pair score a hit or a miss? Second, was the subject's second call, the one involved in the second trial, like the first one (double) or different from it

Table 3

CALL DOUBLES AND NON-DOUBLES IN NON-OVERLAPPING PAIRS OF TRIALS IN RELATION TO HITS AND MISSES ON THE SECOND MEMBER

	Hit on Second Trial of Pair	Miss on Second Trial of Pair	Total
Pairs with Call Doubles	(*a*) 481 (462.7)	(*b*) 1226 (1244.3)	1707
Pairs with Call Non-Doubles	(*c*) 3878 (3896.3)	(*d*) 10495 (10476.7)	14373
Total	4359	11721	16080
Ratio of Doubles to Total	11.0%	10.5%	10.6%

$\chi^2 = 1.11$ (1 d.f.) $P = .29$

(non-double)? In Table 3, all the data have been pooled in a four-fold contingency table to show the over-all relationship between these two variables. The chi-square value of 1.11 obtained from the pooled data is quite insignificant. At first sight, therefore, it would seem that the subject called as many doubles when making misses as when making hits in spite of the fact that her over-all scoring on these trials was highly significant.

Since this analysis did not verify our predictions, we examined the results more closely to see if there might be a genuine effect which was somehow obscured and cancelled out by the pooling of the subject's scores over a period of four years. The figures at the bottom of Table 3 show that the subject called only 11.0 percent of doubles in those pairs in which the second trial was a hit, and 10.5 percent in those pairs in which the second trial was a miss on the target. This shows an over-all strong aversion to calling doubles, since the expected frequency in a random series is 20 percent. It is surprising, therefore, that the subject was able to score, as she did, more than 27 percent hits on the second member of the pairs without approximating more closely to randomness on her number of doubles.

We decided, therefore, to break down the data of Table 3 into the eight approximately equal chronological samples which were used

in our earlier report. When this was done, we found that there was a marked change in the subject's tendency to call doubles between the first and last sections of her total data. During approximately the first two-thirds of her work the subject called fewer than her average number of doubles for the entire series; and during the last fourth (that is, from about the eighty-fifth session to the close of the one hundred and twentieth, which is as far as the present analyses go), she actually called more doubles than the number expected in a random series. Table 4 shows the data of the eight separate samples for the doubling tendency in relation to calls making a hit or a miss and the chi-square evaluation for each sample. This analysis, which exhibits more clearly the variations in the subject's doubling habits throughout the series, shows that the data of three of the chronological samples each gave a significant value of chi square. In each case, the significance is due to the fact that the number of doubles formed by those pairs in which the second call made a hit came closer to the expected number of 20 percent. A

Table 4

CHRONOLOGICAL DISTRIBUTION WITHIN EXPERIMENTAL SERIES OF CALL DOUBLES IN RELATION TO HITS AND MISSES ON THE SECOND TRIAL OF NON-OVERLAPPING PAIRS

Sample Number	Data Sheets	Pairs with Hits on Second Trial			Pairs with Misses on Second Trial			Chi Square* (1 d.f.)	P
		Pairs $(a+c)$†	Call Doubles (a)†	Percent	Pairs $(b+d)$†	Call Doubles (b)†	Percent		
1	88	612	51	8.3	1500	94	6.3	*2.90*	.089
2	76	508	48	9.4	1316	115	8.7	*.23*	.81
3	76	509	42	8.3	1315	98	7.5	*.33*	.78
4	85	561	58	10.3	1479	91	6.2	*10.53*	.0012
5	85	583	47	8.1	1457	67	4.6	*9.47*	.0021
6	85	472	36	7.6	1568	120	7.7	.00	.99+
7	85	529	86	16.3	1511	243	16.1	*.01*	.96
8	90	585	113	19.3	1575	398	25.3	8.37	.0039

$\Sigma\chi^2 = 31.84$ (8 d.f.)
$P = .0002$

*The values of chi square were obtained by the formula appropriate for a fourfold table, as shown in Table 1.

†The letters in parentheses at the top of the columns of this table identify the figures of each sample in relation to the lettering of the cells shown in Table 3. The figures for cells *c* and *d* in each fourfold table may be obtained by subtraction.

better test of the effect for the total data than that given in Table 3, one that takes account of the subject's change in doubling tendency part way through the series, is to sum the chi squares of the eight individual samples in Table 4. This gives a total value of chi square = 31.84 with eight degrees of freedom, which is highly significant (P = .0002).

Discussion

It is interesting to observe how consistently the subject tended to call more doubles in making a hit in the early part of her data in which she was underdoubling, and fewer doubles when scoring a hit toward the end of the series, when she was showing a general tendency toward overdoubling. The first six samples of Table 4 cover the period during which the lowest percentage of doubles was called. The seventh sample represents the transition phase; during the first part of this period of testing, the subject was still calling a low number of doubles, but during the latter part she had risen to more than the chance number. The last group represents data in which she was consistently calling above 20 percent of doubles. In Table 4 the chi-square values in italics are those for the sections of the data in which there is a larger ratio of doubles to non-doubles in the pairs of which the second call scored a hit than in those with the second trial giving a miss. These italicized values predominate during the first part of the series as we would expect, and the only instance in which we have a significant value of chi square that is not italicized is in the last chronological sample.

The general indication from the chronological study, therefore, is that the subject did break away from the doubling tendency which was habitual to her in order to make the calls that scored her ESP hits. This break, as we expected, brought the frequency of doubles in the calls closer to the random frequency that prevailed in the target series. The really surprising fact, however, is that the difference in the numbers of call doubles associated with hits and misses in the target pairs recorded in Table 6 was not even greater than it was found to be. Four of the eight chronological samples showed virtually no difference. It is conceivable, if not probable, that the doubling tendency is the strongest associative pattern with which the

ESP process has to contend. This would make it all the more important to study the other call sequences in relation to hits.

The interpretation of the findings suggested in the preceding paragraph may seem to be the obvious one. However, before it can be offered as the correct psychological explanation of the results, it is necessary to consider another way in which the greater proportion of call doubles with the second member making a direct hit may have been produced. If Mrs. Stewart showed a general tendency to make extra-chance scores on forward displacement as well as to do her high scoring on the direct target, this would have produced the doubles effect under consideration. Each time that a forward displacement hit and a direct target hit occurred on the same target we would have an instance of the special type of call doubles that has been found to occur to a significant degree especially in the fourth and fifth chronological samples. Given the high rate of scoring on the direct target, if forward displacement hits should also be occurring with above-chance frequency, then the times when the two kinds occurred coincidentally on the same target would be higher than expected for a series without positive-scoring forward displacement.

It was necessary, therefore, to examine Mrs. Stewart's data for evidence of forward displacement and to do this in those parts of her records where any genuine tendency to hit the target one step ahead by ESP could be tested independently of her direct-target scoring and secondary effects related to it. We therefore examined those pairs of trials which gave two direct-target misses to see whether there was any tendency for the first call to hit the second target of the pair. To make this check as free as possible of all those factors and peculiarities in her records that might affect displacement scoring, we also left out of account the pairs of misses with a double in either the calls or the targets.

To our surprise, this analysis did yield evidence of forward displacement at a high level of statistical significance. The detailed results will be given in a later report. It does not follow as an inescapable conclusion, however, that the presence of a tendency for forward displacement in the miss areas of Mrs. Stewart's columns explains the excess number of call doubles of which the second member gave a direct hit. The former effect *could* account for the latter;

but whether in fact it *does* do so can only be ascertained by seeing how the two findings are related.

A picture of the relation between the two findings adequate to show whether they are *one* effect or *two* may be obtained by seeing how the two sets of observations compare in significance through the eight chronological sections of the data. If they are both part of one general displacement effect, the results should vary together in significance; whereas if they do not vary together, they would clearly be different effects. The forward displacement in the miss areas was evaluated by a chi-square test in each of the eight sections, and this set of chi-square values was compared with that for the call-doubles data shown in Table 4. The fourth section of data, which gave the highest chi square (10.53, with one degree of freedom) in Table 4, gave the lowest (1.0, with one degree of freedom) for displacement, whereas the sixth section gave the highest for displacement (10.9) and the lowest in Table 4 (.01). A regression test of the relation between the two sets of chi-square values gave a regression coefficient with a negative value, showing that the two sets of observations are clearly not based upon a general forward displacement effect.

This is not to say, of course, that the observed excess number of doubles with the second call a direct hit may not be described as a kind of positive-scoring forward displacement, with the displacement and direct scoring centered on the same targets. But this interpretation would require an assumption that two high-scoring forward displacement tendencies were working concurrently but independently in these ESP tests. We prefer what seems to us the simpler explanation that the subject overcame to a degree her tendency against calling doubles in order to score some of her hits on the direct target.

IV. Frequency of Calling Doubles Following a Hit or a Miss

Problem and Review

The question to be discussed in this section is the converse of that with which we have just dealt. It is this: Does making an ESP hit have any effect upon the subject's associative tendencies which may be revealed statistically by a difference between the number of call symbols repeated after hits and the number repeated after misses?

This problem is one that has already been investigated in other ESP data under the general heading of change of call after ESP success. The question was first raised by Soal (3) in connection with the 1936 work of his two subjects, B.S. and Mrs. Stewart, and the same analysis was applied again (4) to the work with B.S. for 1942-43 as reported by Soal and Goldney. In the case of B.S. particularly it appeared on the basis of these first analyses as if the subject, who was scoring positively on forward displacement, was exhibiting a tendency to change his call after the hits on the cards one step ahead in order to miss these same symbols on the next trial when they were the direct targets. In a further examination of the change-of-call question, however, Pratt (1) pointed out that a general tendency to score negatively on the direct target, coupled with a general tendency to score positively on forward displacement, would produce the appearance of change of call after hits as a statistical artifact. When the B.S. data were examined for such a general negative scoring tendency on the direct target, the effect was found, and therefore the interpretation that this subject exhibited a psychological tendency to change his calls after hits was unsupported.

However, two subjects of the Martin and Stribic experiments produced evidence that a hit had some effect upon the following call that could not be explained in the same way (1). This evidence, while it was not so strong as to lead to any final conclusion that a change-of-call effect had been established, strongly suggested an influence of ESP success upon the call sequence. The importance of this suggestion made it all the more necessary to investigate the problem further whenever an opportunity arose.

Results

We have checked Mrs. Stewart's data, then, for a call sequence effect following hits and misses. In this analysis the 16,080 non-overlapping pairs were again first evaluated as a whole. For this purpose they were separated into two groups according to whether the subject made a hit or a miss on the *first* target of a pair. Each of these two groups was then divided according to whether the second call of each pair was the same as the first or different from it. We would expect the same ratio of call doubles to non-doubles when there was a hit on the first target as when there was a miss if the

Table 5

CALL DOUBLES AND NON-DOUBLES IN NON-OVERLAPPPING PAIRS OF TRIALS IN RELATION TO HITS AND MISSES ON THE FIRST MEMBER

	Hit on First Trial of Pair	Miss on First Trial of Pair	Total
Pairs with Call Doubles (No Change of Call)	(*a*) 451 (424.0)	(*b*) 1256 (1283.0)	1707
Pairs with Call Non-Double (Change of Call)	(*c*) 3543 (3570.0)	(*d*) 10830 (10803.0)	14373
Total	3994	12086	16080
Ratio of Doubles to Total	11.3%	10.4%	

$\chi^2 = 2.56$ (1 d.f.) $P = .11$

two situations are independent. The analysis of the data as a whole, shown in Table 5, reveals that there was no significant difference between these ratios.

As in the previous section, however, the question arises whether there may not have been a chronological variation in effect which is obscured when the data for the four years' work are evaluated as a whole. To test whether this may have been the case, we evaluated the change of call following hits and misses for the eight separate chronological samples used previously. From this analysis, which is summarized in Table 6, we see that there were remarkable variations in the strength of the doubling tendency following hits and misses from one part of the experimental series to another. (The chi-square values are italicized for those sections of the data that showed a larger proportion of calls repeated after hits than after misses.) When the chi squares for the eight separate samples are combined, a total value of chi square = 27.81 with eight degrees of freedom is obtained. This is highly significant, with P less than .0007. We may safely conclude, therefore, that there are some factors present in the different sections of the experiment which are producing a genuine difference in the proportion of doubles following hits and misses, and we are justified in making whatever further study of the data may be required to see how this result may be interpreted.

Table 6

CHRONOLOGICAL DISTRIBUTION WITHIN EXPERIMENTAL SERIES OF CALL DOUBLES IN RELATION TO HITS AND MISSES ON THE FIRST TRIAL OF NON-OVERLAPPING PAIRS

Sample Number	Data Sheets	Pairs with Hits on First Trial			Pairs with Misses on First Trial			Chi Square* (1 d.f.)	P
		Pairs $(a+c)$†	Call Doubles (a)†	Percent	Pairs $(b+d)$†	Call Doubles (b)†	Percent		
1	88	538	53	9.9	1574	92	5.8	*10.06*	.0015
2	76	517	45	8.7	1307	118	9.0	.05	.91
3	76	445	27	6.1	1379	113	8.2	2.15	.14
4	85	476	26	5.5	1564	123	7.9	3.11	.078
5	85	494	28	5.7	1546	86	5.6	*.01*	.96
6	85	487	53	10.9	1553	103	6.6	*9.48*	.0021
7	85	461	86	18.7	1579	243	15.4	*2.81*	.093
8	90	576	133	23.1	1584	378	23.9	.14	.85

$\Sigma\chi^2 = 27.81$ (8 d.f.)
$P < .0007$

*The values of chi square were obtained by the formula appropriate for a fourfold table, as shown in Table 1.

†The letters in parentheses at the top of the columns of this table identify the figures of each sample in relation to the lettering of the cells shown in Table 5. The figures for cells *c* and *d* in each fourfold table may be obtained by subtraction.

Discussion

The chief characteristic of Mrs. Stewart's ESP performance, of course, is her outstanding positive score on the direct target, the one on which the sender was focusing at the time she made her guess. The two most significant chi squares in Table 6—those for samples 1 and 6—indicate that the subject repeated her call more frequently following a direct hit than she did after a miss. Now this effect would arise as a statistical artifact if the subject made a positive score on backward displacement in these trials, just as the finding considered in the preceding section might have been an artifact of forward displacement. A moment's reflection will show how backward displacement would give the appearance of repeating a call after a hit. We know that Mrs. Stewart showed positive scoring on the direct target. If in addition backward displacement to the target of the immediately preceding trial also occurred in the series, and if hits of these two kinds were made independently of one another in a statistical sense, we would occasionally expect to find both forms

of ESP operating in relation to a particular target. For this to happen it would be necessary for the subject to give the same call on adjacent trials, as it is possible for the direct call and the following call both to hit the same target only if they are identical. Following a miss on the direct target, however, the tendency of the subject to score a hit on backward displacement would actually reduce the number of call doubles, as the subject would necessarily have to change her response in order to make a backward displacement hit after a miss.

To see whether a *general* backward displacement tendency was the explanation of the change-of-call finding, we analyzed Mrs. Stewart's data for backward displacement in those pairs in which neither trial gave a direct target hit. We also omitted pairs with call or card doubles, as we did previously for forward displacement. There was some suggestion of positive scoring on backward displacement, though not at so high a level of significance as for the forward displacement. These details also belong in a later report. To see whether displacement may explain the excess of call doubles following hits, we may compare the significance of this analysis with that for the change-of-call results in Table 6, confining our attention to chronological samples 1 and 6, which gave the most striking evidence of ESP influence upon the doubling tendency. The results showed that there was no general backward displacement tendency in those two sections of Mrs. Stewart's work, as one gave a chi square of .4 (one degree of freedom) and the other a value even closer to zero. Consequently, the significantly greater proportion of double calls following hits could not be explained as an artifact due to general positive scoring on backward displacement. We conclude, therefore, that there is a real difference in the sequence of calls following hits and misses in Mrs. Stewart's data and that the earlier, suggestive evidence from the Martin and Stribic series has been confirmed, though here we have a repeat instead of a change of call after a hit.

What psychological interpretation is it possible to place on this finding? In the earliest investigation of change of call, it was thought that such a result might provide experimental evidence that a subject has some kind or degree of knowledge of his own ESP success, inasmuch as he appears to be reacting to it. In the

later discussion of the results, however, it was recognized that there was an alternative interpretation which could not be completely ruled out. This was that certain targets in the series somehow became effective for the ESP process and that the subject sometimes reacted to one of these targets, not by one call only, but by two—or even more—of his calls. From the earlier evidence, it appears that the characteristic reaction was to change the call after success. On this "spread-of-ESP-influence" view, we may suppose that in those cases a subject scored positively on such an "effective" target when it was the direct target and then responded to it by psi-missing on his following call, thus producing the change of call after success. In the present case, Mrs. Stewart tended to repeat the call symbol that scored a hit. If this is an effect of a particular target lasting through the two trials, it becomes a matter of actually *hitting* such a target for two trials, which requires calling a double.

We still see no way to distinguish finally between these two possible interpretations of the change-of-call effect: (a) a reaction to ESP success reflected, perhaps subconsciously, by a change or repeat of call; and (b) a spread of influence around certain targets that somehow influence more than one ESP response. It may be, however, that there is less difference between these two ways of looking at the results than might at first appear. When we say that a change in a subject's patterning of his calls is an indication of reaction to success, we tend to think in terms of a conscious awareness of the ESP hits. On the other hand, we must recognize that there is no sharp division between consciousness and unconsciousness, but that these concepts are abstractions representing the two extremes of a continuum. Parapsychologists have been forced by their failure to find experimental evidence to the contrary to assign psi processes to the unconscious end of the range of mental activities. The results of psi tests probably represent the most conclusive evidence available that unconscious mental processes do occur. The research workers dealing with the psi effects have not ceased to be interested, however, in finding some special situation or device which will enable the ESP subject to show that he is sometimes influenced, be it in ever so subtle or momentary a manner, by the fact that he is employing his ESP faculty. For example, Mrs. Stewart was apparently influenced by

some aspect of the situation when she scored a hit, and this influence was strong enough to cause her to repeat the call and thus score a second hit on the same target. This is a step, although a small one, toward our objective, the discovery of some method which enables the subject to signal when he has successfully used his ESP ability.

V. Summary

In previous analyses made upon the ESP test records of Mrs. Stewart, we found that she scored significantly above chance expectation when the target which she was trying to hit was unlike the following target. However, she demonstrated a significant degree of psi-missing when the target was repeated on the following trial. The basic analysis which led to this discovery was made upon test data extending over an interval of more than four years. A total of 1,340 columns of 25 trials each were involved, which provided 16,080 non-overlapping pairs of trials for the investigation. The tendency for the scoring rate on the first of target doubles to be lower than that on the first of target non-doubles was found to be present to a significant degree in every one of eight chronological sections of the data.

In the present study, these same pairs of trials have been analyzed for scoring effects related to *call* sequences. Three questions have been considered.

1. Is there any difference in the tendency to score low on the first of target doubles in those pairs with call doubles and those without them? The analysis carried out along these lines showed that the below-chance scoring reported in the earlier paper was confined to those pairs in which the second call was unlike the first one. When the subject repeated her call, the first trial of the pair scored just as high on the first of a target double as on the first of a target non-double. This finding suggests the interpretation that the subject did not always react to a target double *as such*, and thereby bring into play her peculiar negative scoring effect which was ordinarily present on sequences of identical targets. When she failed to appreciate the fact that she was dealing with a sequence of identical targets, or whatever it was about the situation that ordinarily made her show psi-missing, she exhibited her usual positive scoring tend-

ency on each symbol separately. The operation of a positive scoring tendency would, in the absence of the effect of avoiding target doubles, lead to more than the expected number of the call doubles being correct. This is what the findings show, and we therefore are led to the conclusion that the negative scoring effect was not present on all of the target doubles.

2. Is the frequency of calling doubles in the paired trials affected by the scoring of hits on the second trial? The analysis of the data as a whole as they bear upon this question gave insignificant results. This showed that if there was any such effect in the data as a whole, a tendency to form *fewer* doubles by the calls making hits in some parts of the series was cancelled out by a tendency to form *more* doubles when making hits in other sections. When the analysis was made on the basis of the eight chronological sections of the data, we found that this cancellation had indeed taken place. Significant results were found which indicate that the number of doubles in the calls tends to approach more closely to the expected chance number (as they were in the target series) when the second call of the pair was a hit. In other words, the subject, in scoring hits in the first part of the long series, overcame her tendency to call too few doubles, and in the last part, her tendency to call too many of them. The evidence indicates, however, that her ESP processes in the first part of the series did not completely overcome her habitual aversion to calling doubles. This general call-sequence effect was not explained by any general tendency to score *positively* on forward displacement.

3. Is there any difference in the tendency to repeat on the following trial a call that makes a hit and one that scores a miss? This question was, like the previous one, answered in the affirmative, and in much the same way: the results were not significant in the data as a whole, but the analysis in terms of the chronological sections showed that there were significant change-of-call effects which fluctuated widely throughout the five-year testing period. Two of the eight sections, the first and sixth, were of outstanding significance, the effect being one of *repeating* the call more frequently after hits than after misses. Again, this finding cannot be explained in terms of a general displacement tendency as there was no evidence of positive backward displacement apart from those trials which followed

hits on the direct target. No way is seen of distinguishing at this time between two possible interpretations of this finding: (a) that the subject was reacting to the fact that she had scored an ESP hit; (b) that occasionally a target affected not only the call of its own trial but the following call as well.

The analyses of Mrs. Stewart's records reported thus far have already shown that, for her, the ESP process was a tremendously complicated network of tendencies working at cross purposes with one another. And there are more findings to come, chiefly dealing with position effects on scoring within the two columns of the page. Significant as her total deviation in the experiments was, we can now see that it represented only a net ESP result, one which fell far short of giving a true and complete picture of the psi processes in her test performances.

References

1. Pratt, J. G. Change of call in ESP tests. *J. Parapsychol.*, 1949, **13**, 225-46.
2. Rhine, J. B. The problem of psi-missing. *J. Parapsychol.*, 1952, **16**, 90-129.
3. Soal, S. G. Fresh light on card-guessing—some new effects. *Proc. Soc. psych. Res.*, 1940, **46**, 152-98.
4. ———. A note on negative deviation. *J. Parapsychol.*, 1944, **8**, 311-15.
5. Soal, S. G., and Pratt, J. G. ESP performance and target sequence. *J. Parapsychol.*, 1951, **15**, 192-215.

Parapsychology Laboratory
Duke University
Durham North Carolina

Queen Mary College
London University
London, England

WHY THROW DICE?

By R. A. McConnell

The minds of layman and scientist alike leap to more exciting and respectable methods for detecting psychokinesis, to methods such as the disturbance of balanced mechanical systems, the triggering of chemical reactions, the levitation of objects, the control of living organisms, or the influence of radioactive decay. That the experimenter does not at once begin research upon these and other proposals is likely to be taken as indicating a lack of imagination.

The question "Why throw dice?" is usually completed by the words "instead of doing X," where X is offered as an ostensibly more attractive research procedure. It will illuminate this question to let X be a specific proposal whose merits can be examined. The comments that follow derive from one particular view of the research situation. Their only claimed value will be to show that there are diverse considerations in planning a program in this field.

The most frequently suggested way of detecting psychokinesis is to use a balanced mechanical system. It is often said, "If psychokinesis is real, it ought to be possible to deflect a delicately balanced needle." So phrased, this is a presumptuous statement. No one can say that the phenomena of nature "ought" to have certain characteristics. If our intuition could be trusted, psychokinesis ought not to occur at all.

On the other hand, it is permissible to speculate as to whether one might expect a balanced needle to exhibit psychokinesis. How does a balanced needle compare with a falling die? It is conceivable that the "emergence" of a single quantum of (visual) energy applied in the optimum way at the optimum instant could shift a tumbling die from one "causal track" to another. But to deflect perceptibly a balanced needle would require many such quanta. While this kind of argument has no experimental justification, it at least suggests that deflection of a balanced needle could be a gross effect that might be possible only with exceptionally gifted persons, or might not be possible at all.

The susceptibility of a balanced needle to movement by psycho-

kinesis cannot be judged *a priori*. But the suitability of this sort of device from a psychological point of view can be assessed with some confidence. Imagine the attitude of a student volunteer who has been introduced to this stationary device and asked to make it move. If the student has been treated honestly, he will perceive by his own mechanical intuition that the needle could not move by itself. It will take a highly persuasive experimenter to overcome the self-conscious incredulity of the student under this circumstances.

Compare this situation with that encountered in handling dice. If the dice are falling they are bound to turn up one face or another. The intrinsic plausibility of a mental influence is for most people far greater for moving dice than for a stationary needle. It is well known that many persons have a positive belief that dice can be controlled by direct action of the will. Moreover, dice are usually associated with ideas of play. This makes it easy to create a game-like atmosphere in the laboratory.

Past research indicates that the *consciously controlled* evocation of psi demands a relaxed, earnest, and believing attitude. The fact that spontaneous psi frequently occurs in quite different emotional situations seems to have no bearing upon the ordinary laboratory experiment (an idea that is well worth the attention of the would-be investigator). If any lesson regarding the balanced needle can be drawn from the literature of psi it is probably this: the psychological conditions in a balanced needle experiment would be unfavorable for the appearance of psychokinesis as compared with the conditions in a falling-dice experiment.

Let it be supposed, however, that a self-confident and favorably disposed "subject" sits down before a balanced needle. Gross psychokinetic phenomena, if they occur at all, do not occur upon demand. It is reasonable to expect that a long experimental session might yield only one detectable movement of the needle, if any. When confronted with an occasional unpredicted deflection of a mechanically balanced system, the skeptical physicist will justifiably ask: "How do you know that the needle was so dynamically balanced or de-coupled as to eliminate the possibility of resonance with natural or man-made earth motions?" To validate even tentatively a causal relationship in such physically ambiguous circumstances

would require statistical treatment, which in turn implies a reasonable number of successes. Until a subject appears who claims the requisite ability, it would appear of dubious value to plan for such an experiment.

Consider how this situation compares with dice. Except for rare instances of cocking, the dice offer the certainty of one from a mutually exclusive set of events. One out of the six faces must come up. Thus, the dice offer an ideal method of control. Each one of the mutually exclusive events (faces) can be made the target in successive groups of trials. Whatever bias may exist in favor of one target will be offset by the opposing bias that must exist for the other events. The total probability of all possible events is *one*, or certainty. Exactly how the component probabilities are distributed among the faces is of no consequence for the detection and measurement of the psychokinetic effect.

Before discussing the broader questions that might guide a psychokinesis research program, it may be useful to consider another specific research proposal. It has often been said that a person with psi capacity might be able to influence radioactive decay. The disintegration of a nuclear particle is believed to be a wholly random event which cannot be specified except statistically, as is usually done by stating the "half-life" for the reaction in question.

Some of the difficulties in the balanced needle experiment exist also for the control of radioactivity. But these might be regarded as offset by the scientific attractiveness of the proposal. The discovery of a way to affect radioactivity would point toward practical applications in nuclear physics. Such a discovery would also be of great theoretical interest. The psychokinetic interaction between mind and inanimate matter has been experimentally demonstrated only for molecular aggregations. The proof of interactions at a nuclear level would affect the nature of any theory that attempts to unify psi phenomena with the rest of physical knowledge.

From the physicist's point of view, a mental influence upon nuclear behavior is more plausible than an influence upon dice. Up to now there has been no satisfactory theory to explain what happens inside the nucleus. There is, moreover, no known way by which the disintegration of a radioactive nucleus can be predicted,

much less speeded up or delayed. Here is an area of ignorance in which anything could happen. It is the kind of area in which strange events could be most easily accepted.

Granted that the psychokinetic control of nuclear decay is an attractive area for investigation, what are the drawbacks? A prominent feature of this approach would be its cost: small in terms of what is usual in nuclear research, but large compared with the sums now available for psychokinetic research. The instrumental complications in setting up statistical controls suitable for the detection of marginal influences are not encountered in most nuclear research, but they can be easily appreciated by the worker in that field.

Let it be supposed that the nuclear apparatus is built and ready to go. How does one instruct the subject so that he will "exert an influence"? There is a gap in faith and in feeling between this experiment and dice throwing that may be too large for direct bridging. The difficulty in all psi research is that when the phenomena are completely absent, no conclusions, positive or negative, are warranted except in so far as the new experimental situation can be related closely to the old. Our knowledge of psychokinesis in the dice situation is so fragmentary that no one can guarantee to produce the effect on demand. Under these circumstances, how much of a gamble in a distinctly different kind of experiment is warranted?

The answer lies in the over-all world program in psi research. As long as professional investigation remains a small-scale operation by a handful of scientists, the work must be pitched at a corresponding level. The kind of free-shooting that is commonplace in nuclear physics is for the present a practical impossibility in this field.

This can be explained in a different way. Many workers believe that the most promising approach toward greater knowledge about these phenomena is to find an experiment that is "repeatable by prescription." There has already been enough partial success to warrant continuing attempts in this direction. With a repeatable experiment available, the systematic investigation customary in other fields of science could begin. If this objective is sound, it bears upon methodology in the following way. The inhibiting factors which prevent the appearance of psi phenomena and thereby obstruct experimental repetition are *psychological and not instrumental.*

This is not an absolute statement. It could prove to be wrong. But it is strongly indicated by what we already know. Its practical implication is that cautious self-discipline must be exercised in setting out upon new instrumental approaches. The kind of gamble that will be warranted will depend upon the stake that one can afford to lose.

These considerations suggest that dice-throwing will have a continuing important place in whatever program may be undertaken in psychokinesis. The relative weight to be given dice work will depend upon the magnitude of the program as a whole. As already discussed, the favoring features of dice are three: 1. They work. Psychokinesis has been demonstrated by this method. 2. They are psychologically easy to use. Dice-throwing elicits a playful, believing attitude in the subject. 3. They are well suited to statistical control. Falling dice provide an experimentally simple set of mutually exclusive events having nearly constant probability.

Among the questions that may be answerable by the use of dice are these: How dependably would psychokinesis appear if college students could begin an experiment casually, without prior scheduling? Is the presence of an experimenter an aiding or inhibiting factor for most people? How do measurable personality attributes correlate with psychokinetic ability? What about physical environment? What happens when more than one subject at a time tries to affect the dice? Would uniform instructions to the subject by means of tape-recording lead to more uniform psychokinetic performance? Can dice be influenced from great distances? Can a subject influence dice while he is sleeping naturally or hypnotically?

Partial answers have already been obtained to some of these questions within the limits of available dice-throwing equipment. Some of these answers have been published; others are awaiting publication. Considerable information is already available about corresponding questions in ESP experiments. It will be interesting to see whether the results for ESP agree with those for psychokinesis.

Department of Physics
University of Pittsburgh
Pittsburgh 13, Pennsylvania

MINOR ARTICLES

EFFECT OF NOVELTY IN TEST CONDITIONS ON ESP PERFORMANCE

By Remi J. Cadoret

ABSTRACT: The author, using himself as subject, went through 272 runs of ESP card tests trying to induce high scoring by the introduction of certain novel conditions during the test. He tried such conditions as reading poetry or singing aloud during the card runs. High scoring rates were made in those trials in which novel test conditions were used for the first time. Negative deviations resulted when these conditions were no longer novel—that is, when they were used a second or more times. The difference between these "new system" and "old system" runs is a significant one.

The number of trials in the run with any particular condition varied according to the mood of the subject. Runs containing 14-25 trials gave a decline in scoring which showed a significant difference between the first and second halves of the run; whereas shorter length runs (1-13 trials) showed a slight incline in scoring.

Mr. Cadoret is a third year medical student at Yale. He has contributed one other article to the Journal, a collaborative report with Dr. J. G. Pratt on the consistent missing effect in ESP.—Ed.

Introduction

To produce high scores at will has been the goal of many a researcher in ESP. The approaches have varied from studies of ESP in the "spontaneous" cases and experiments in which the influence of hypnosis, drugs, and other conditions were tried, to minute study of ESP-personality correlates. While each of these approaches has borne some fruit, we still possess no sure method of releasing ESP from the depths of the unconscious.

Subjects as well as experimenters have sought various psychological conditions or "systems" which they felt would enable them to perform better in tests. Some have felt that singing aloud, or listening to music, or holding the card deck in the hands while calling the cards is conducive to higher scores. Such a system has recently been tried in a PK experiment by Thouless (**7**), who found that reciting poetry aloud during his self-testing sessions raised his score significantly over what it had been just prior to the introduction of the poetry. He noticed, however, that there was a decline in the efficacy of the system after it had been in use for a short time. The

explanation of the higher scoring appeared to lie, not in the system as such, but in its novelty effect.

Researchers in ESP have also obtained a rise in scoring through the use of novel stimulus material. This can be seen in the Pratt-Woodruff experiment (**1**) in which symbols of various sizes were used. The experimenters concluded that "when the results were considered in relation to the amount of the subject's experience with the different kinds (sizes) of stimulus material, subjects were found to score significantly better with the new than with the old material. The advantage of working with new material was found to decline after a time as the subjects became more experienced with the new material." The experimenters further noted that if the subjects were unaware that new symbol sizes were being introduced they did not do better.

Another remarkable example of the effect of novelty is Zirkle's experiment performed with Hubert Pearce after the latter had apparently lost his ability to score high (**4**). The subject was given 272 runs under 21 different experimental conditions such as "staring at a fixed point, looking into a crystal ball, following a rhythmic tapping, or walking about the room." As soon as the subject tired of one condition he was switched to the next. When the scores made under each condition were divided chronologically into two halves it was found that the average for the first half (5.44) was significantly higher than the average for the second half (4.90).

The similar injection of novelty into an otherwise routine experimental situation has also been observed by Rhine (**2**) and Smith and Gibson (**6**) to have a favorable effect on scoring.

The present paper reports an experiment in which I was both experimenter and subject. In it I tried to induce a high rate of scoring by the use of various systems, changing often to obviate any decline phenomenon.

General Procedure

I did all the runs at home as a relaxation from medical studies. The DT technique was used throughout, with the card deck being shuffled at least five times (dovetail shuffle), cut, and placed out of sight. This procedure was performed with my gaze averted or eyes closed. I recorded the calls on the standard forms used by the Parapsychology Laboratory.

As for the systems, I tried any procedure that particularly struck my fancy at the moment. Some of the systems I used were: reciting poetry, both aloud and to myself; reading poetry; reading and reciting prose; singing aloud; studying; whistling; writing the calls on the sheet in haphazard order, or writing them "up through" while checking "down through." In short, I tried to use any system that kept my mind off the task and the goal of scoring high, with all its attendant anxieties. Some systems I used once, others often, depending on my whims at the time.

By way of a "control," whenever I was not inspired with an idea for a new system or did not feel like using an old, well-tried one, I merely wrote my calls DT without any conscious effort at employing a system. These control runs or parts of runs were liberally interspersed with the sequences of system trials, the order of control and system periods being subjectively determined on each occasion. For convenience, the term "run" in this paper will refer to any group of control or system trials, from 1 to 25 in number, done at one time. Any one column of 25 trials might contain both control and system trials. For example, there might be seven trials by one system, followed by eight control trials, and ten trials by another system. On no occasion was a portion of a column checked by itself (since a closed deck was used); I waited until all 25 cards in the deck had been called. During Series 1 the points where runs of control or system trials started were marked on the record sheets as the calls were made, with brief notations of the systems used.

No definite number of runs was set for any series nor for the experiment as a whole, since I planned to keep going only as long as new systems suggested themselves to me without much effort. The end of a series was determined by my stopping for a period of a few days to a week by way of a respite. Only then did I check all the results to date in a way that showed the effect of the systems. While this method is highly arbitrary it enabled me to work without a feeling of pressure that would inevitably be present if there were a set number of runs, both system and control, to be completed in each series. This was to be a simple pioneering investigation; the niceties and restraints of rigid control of conditions, I felt, would have a deleterious effect on the all-essential enthusiasm, and could be added later.

Results

Four series were carried out. Series 1 lasted from May 14 to 18, 1951. My idea at this time was simply to see whether there was any difference[1] between system and control trials. The over-all scores are shown as follows:

System Runs					Control Runs				
Trials	Hits	Dev.	CR	P	Trials	Hits	Dev.	CR	P
758	176	+24.4	2.22	.03	442	94	+5.6	0.67	.5

$\chi^2 = .61$

While the deviation of the system runs is almost significant, it is not significantly greater than the control run score. During this series I formed the impression that new systems worked better than those that had been in use several times. To test for this chronological decline proved difficult since one system might be used once or twice a session, another system many times, still others only once during the entire experiment. It seemed best to divide all the system runs into two categories: (1) Trials obtained the very first time a new system was used anywhere in the entire experiment, whether a complete run of 25 calls or only part of a run. To be considered "new" a system had to be either one that had never been tried before in the entire experiment or a new combination of two or more old systems whose product represented a novel condition. (2) Calls made with old systems; that is, anything done after the initial use of the system. By dividing the system runs in this way all the data could be used and any new-system salience would become evident.

This method of dividing the data of Series 1 gave a high rate of scoring for the new system runs and a significant difference between the new and old system runs (see the first section of Table 1).

In spite of the fact that the division into new and old systems in Series 1, which had to be done after the results had been checked, was carried out in as objective a manner as possible, there is still an outside possibility that *post facto* selection may have played a part. So to test the apparent difference between new and old system runs

[1] The comparisons of the rates of scoring under any two conditions presented in this paper were made in each instance by forming a fourfold table of the observed number of hits and misses found in each of the two conditions. This fourfold table was then evaluated by the usual chi-square method, which gives a value of ^{2} with one degree of freedom.

Table 1

OVER-ALL SCORES OF SYSTEM AND CONTROL RUNS IN SERIES 1-4

Condition	Trials	Hits	Dev.	CR	P	χ^2
SERIES 1						
New system	270	83	+29.0	4.41	.0001	13.32
Old system	488	93	− 4.6	0.52		
Control	442	94	+ 5.6	0.67		(P=.0003)
SERIES 2						
New system	105	24	+ 3.0	0.73	.01	4.28
Old system	323	46	−18.6	2.59		
Control	347	62	− 7.4	0.99		(P=.04)
SERIES 3						
New system	215	57	+14.0	2.39	.017	3.42
Old system	485	98	+ 1.0	0.12		
Control	450	76	−14.0	1.65		(P=.06)
SERIES 4						
New system	144	42	+13.2	2.75	.006	9.99
Old system	403	68	−12.6	1.57		
Control	598	106	−13.6	1.39		(P=.002)
TOTAL SERIES 2-4*						
New system	464	123	+30.2	3.50	.0005	16.97
Old system	1211	212	−30.2	2.17	.03	
Control	1395	244	−35.0	2.34	.02	(P=.00004)

*Since these three series were performed under identical conditions and for the same purpose, that of seeing whether the observed difference between new and old systems in Series 1 could be confirmed, it is valid to combine their scores.

further, I carried out three more series, with the additional precaution that the fact that a system was *new* was also noted on the record sheet at the time the calls were made and before the targets were seen.

The results of Series 2, 3, and 4 are shown in Table 1. The new system scores are consistently positive. The difference between them and the old system scores is highly significant when the series are combined. Thus the findings in the first series are substantiated: new systems are associated with a higher scoring rate. The extreme

negative deviation of the control runs, however, is in marked contrast to the slightly positive deviation in Series 1.

The distribution of scores within the run has frequently shown significant trends in scoring in ESP research, chiefly a decline in success. Accordingly, a search for decline patterns was made in these data and some interesting differences in decline were noted. Runs (or fractions thereof) were divided into a top and bottom half, with the middle trial being omitted in case of an odd number. The results, given in Table 2, show that all of the positive scoring occurred in the top half of the run. However, the evidence for a decline is only suggestive in both system and control runs.

Table 2

DECLINE PATTERNS IN SYSTEM AND CONTROL RUNS

Condition	Trials	Hits	Dev.	CR	P	χ^2
System runs—Top half.....	947	224	+34.6	2.81	.005	3.80 (P=.05)
System runs—Bottom half....	947	189	− 0.4	.03		
Control runs—Top half.....	859	178	+ 6.2	0.53		4.41 (P=.04)
Control runs—Bottom half...	859	144	−27.8	2.37	.02	

A breakdown of the runs to show new and old systems separately reveals that the decline in the system runs was contributed by the trials done with old systems. (See Table 3.)

Since many of the new system calls were done in much shorter sequences than the standard run of 25 calls, it seemed a logical step

Table 3

DECLINE PATTERNS IN NEW AND OLD SYSTEM RUNS

Condition	Trials	Hits	Dev.	CR	P	χ^2
New system runs—Top half..	278	86	+30.4	4.56	.000005	0.21
New system runs—Bottom half.....	278	81	+25.4	3.81	.0001	
Old system runs—Top half...	669	138	+ 4.2	0.41		4.41 (P=.034)
Old system runs—Bottom half.....	669	108	−25.8	2.49	.013	

to examine the runs of shorter length to see if any peculiarity in their hit distribution might explain the lack of decline in the new system runs. Accordingly the data were divided into runs made up of 1 to 13 trials and another group of runs containing 14 to 25 trials. The division point at 13-14 was purely arbitrary and it was made without reference to the data. The results of dividing the data in this manner appear in Table 4. Runs of 1-13 trials showed no evidence of a decline, but rather, an "incline"; and, by contrast, the longer runs with 14-25 trials demonstrated significant decline patterns both in the system and in the control groups. The possible reason for this will be dealt with in the discussion section.

Table 4

DECLINE PATTERNS IN SHORT (0-13 TRIALS) AND LONG (14-25 TRIALS) RUNS

Condition	Trials	Hits	Dev.	CR	P	χ^2
RUNS WITH 0-13 TRIALS						
System runs—Top half.....	89	23	+ 5.2	1.30		.68
System runs—Bottom half....	89	28	+10.2	2.63	.01	
Control runs—Top half.....	149	33	+ 3.2	0.59		2.13
Control runs—Bottom half...	149	44	+14.2	2.85	.0044	
RUNS WITH 14-25 TRIALS						
System runs—Top half.....	858	201	+29.4	2.48	.013	5.62 (P=.018)
System runs—Bottom half....	858	161	−10.6	0.88		
Control runs—Top half.....	799	168	+ 8.2	0.70		6.60 (P=.01)
Control runs—Bottom half...	799	128	−31.8	2.79	.0053	

DISCUSSION

This experiment offers the following evidence that extrachance factors were in operation: (1) the significant differences between old and new systems in Series 1 and in Series 2-4 combined; (2) the significant positive scoring of the new systems in Series 1 and in Series 2-4 combined; (3) the highly suggestive and in some cases

significant decline patterns which to a considerable degree are consistent with the declines found in many published ESP series. The first two lines of evidence are not statistically independent, but the third is independent of the first two.

Before we can interpret these extrachance factors as ESP we must see whether the experimental conditions are adequate to rule out the counterhypotheses which might be offered as possibly accounting for the results. Sensory cues from the cards are reduced to a minimum by the use of the DT procedure in the entire experiment. The shuffling (minimum of five times dovetail) and cutting are adequate to insure a random target order. Recording, however, was unwitnessed and therefore open to question. Consequently a check was made of the number of times each symbol appeared in each run of twenty-five target cards at which the calls were aimed. The count revealed that each run contained five of each symbol. This pretty well excludes the likelihood that the wrong symbol was recorded in accordance with "wishful thinking."

Since the length of series (or of the experiment for that matter) was not predetermined, the question of optional stopping must be raised and answered. The major finding here has been the difference in scoring between old and new systems. Now this difference was highly significant in Series 1, which was stopped when I had only a general idea that a difference existed and before any effort had been made to examine or evaluate it. Optional stopping cannot explain the fact that the difference was *significant* at the point chosen for ending the series. Moreover, Series 2-4 are also significant in respect to the difference between new and old systems. The stopping-point of each of these series was not determined by the difference between old and new system scores but by other factors such as fatigue. In fact, the data were not analyzed until each series was ended. Hence the end of the series has no relation to the most advantageous point at which to halt.

A further point to consider is the consistency of the findings: in four series the new system runs were always higher than old system runs, significantly so in two of these series. The results conform throughout the entire experiment to the general principle of new systems favoring higher scoring than old systems, not to the fluctuations of a random series which was simply ended at the point

of accidentally reaching a designated level above expectation. It seems rather difficult to invoke coincidence in order to account for such consistent results.

Since the most likely counterhypotheses in this experiment seem to be adequately eliminated, it is most probable that ESP was responsible for the extrachance results reported here.

The declines found in this experiment seem to be associated with the length of the run. It was found that declines were present only in the longer runs of 14 to 25 calls, while in contrast, the pattern in shorter length runs is an incline. There is one essential difference between the shorter runs and longer runs which may explain this. The shorter runs represented calls made when each trial was more or less a distinct unit or, tentatively, a possible stopping-point. On the other hand, the longer runs consisted almost entirely of complete runs of 25 trials which I began with the intention of completing without changing conditions. Therefore the shorter runs represented quite a different task from the others in that I was often uncertain how many calls I would make in them. For this reason the fragment was never as well structured in my mind as would be the standard run of 25 calls. Structuring of the run or task by subjects has been considered one of the possible causes for decline and, more especially, salience patterns (3). Thus the lack of a definite task in the shorter runs may have prevented any declines from appearing in them.

One aspect of the results that is not so easy to explain but which is too important to be overlooked is the psi-missing tendency shown by the old system and control runs. This was not present to a significant degree in Series 1, but in Series 2, 3, and 4 these two conditions together totaled 2,605 trials with a negative deviation of 65.2 hits. This deviation has a CR of 3.19 and, in spite of the fact that the result was not predicted, it probably represents a genuine case of negative-scoring ESP. Both the control and the old system runs, as shown in Table 1, contributed about equally to this below-chance scoring. It is interesting, therefore, to compare the system and control trials with respect to the position effects found. The decline curves for both system and control groups in runs of 14-25 calls are very similar, and both groups show the incline in the runs of 1-13 calls. So it appears as if the same or similar processes were going on concurrently in both system and control groups. There is no ob-

vious explanation for the negative deviation of the control runs. However, most of the negative deviation in the control runs as well as in the old system runs is contributed by the runs of 14-25 trials. This may indicate that short length runs may have been responsible for a good deal of the high scoring observed in the new system runs.

To check this point, all new system runs were divided into those of 1-13 and 14-25 trials. The results of this division demonstrate that while the shorter runs gave a slightly higher average, the longer length runs by virtue of the larger number of trials involved, contributed most of the positive deviation obtained:

Length of Run	Trials	Hits	Dev.	CR	P	x^2
0-13.....	148	43	+13.4	2.75	.006	.09
14-25.....	586	163	+45.8	4.87	.0000012	

The difference in scoring rate between shorter and longer length runs is not significant. Thus it appears that the novel system worked toward producing sustained positive scoring whatever the length of run.

Here we have seen how a large number of different systems were successful at first use in raising the scoring rate significantly. These results agree essentially with the findings of other investigators (**1, 2, 4, 6, 7**) and with many clinical impressions that have been recorded: that psychological novelty in some form improves scoring and that the effect, while often dramatic, is characteristically short-lived. By what psychological mechanism the scoring benefits is still a mystery, although we may still surmise.

For one thing, there seems to be a large element of spontaneity involved in these systems. I use "spontaneity" in the way Scherer (**5**) defined it: a state in which one's choices are made without rational selection. The rational part of one's personality may become so involved and absorbed in carrying out the little details necessary in performing a system properly that it does not have opportunity to "censor" one's calls, if that indeed is the function of this rational part. It is easy to conjecture that the system can function in this way only as long as it absorbs the attention of this rational part or

whatever essential part be involved. A new system seems to be effective in giving high scores in my case for only a very short time, but we can derive some comfort from the knowledge that in others (**4, 7**) the censor seems to have been engaged for longer periods of time.

It might be instructive (though not without danger) to compare the mechanism of ESP as apparently revealed here with the somewhat better-known physiology of vision, especially night vision. Since the rods of the eye are especially sensitive to dim light, the most sensitive portions of the retina are at the periphery where the rods are most numerous. To see an object in poor light, as in the late evening, one must look at the object with averted gaze so as to catch it in the peripheral visual field. Most people look directly at the object unless specially trained and so do not benefit from using their rods. The reader has probably seen where the analogy lies. Macular or central vision may represent that part of my consciousness that focuses on a system while my "peripheral vision" receives ESP impressions.

Various methods have been tried to rid oneself of this rational, inhibiting aspect of the personality, or whatever it may be. These methods have ranged from simple tricks, like systems and other novel psychological features, to the more complex methods employed by ancients and moderns: trance, scrying-glasses, crystals, drugs, and the like. But the fact that in the present study a large number of different systems were capable of raising scoring certainly makes it appear that the essential factor was not the emotions (as when reading poetry), nor the various motor, speech, and other brain "centers" (as when whistling, singing, tapping feet, etc.) but a common denominator to the use of all the systems: the fact that each new system involved some essential part of the subject's consciousness.

References

1. Pratt, J. G., and Woodruff, J. L. Size of stimulus symbols in ESP. *J. Parapsychol.*, 1938, **3**, 121-58.
2. Rhine, J. B. Conditions favoring success in psi tests. *J. Parapsychol.*, 1948, **12**, 58-75.
3. ———. Terminal salience in ESP performance. *J. Parapsychol.*, 1941, **5**, 183-242.

4. RHINE, J. B., *et al. ESP after Sixty Years.* New York: Henry Holt, 1940.

5. SCHERER, W. B. Spontaneity as a factor in ESP. *J. Parapsychol.*, 1948, **12,** 126-47.

6. SMITH, B. M., AND GIBSON, E. P. Conditions affecting ESP performance. *J. Parapsychol.*, 1941, **5,** 58-86.

7. THOULESS, R. H. A report on an experiment in psychokinesis with dice, and a discussion on psychological factors favoring success. *Proc. Soc. psych. Res.*, 1951, **49,** 107-30.

533 Cedar St.
New Haven, Connecticut

A STUDY OF THE EFFECT OF CONDITIONING ON MULTIPLE-ASPECT ESP SCORING

By Stanley J. Hallett

ABSTRACT: Multiple-aspect targets in ESP tests are targets in which the subject must make two or more identifications—such as suit and number—in order to get a complete hit. A preliminary experiment was conducted in which the two target aspects were: (1) the ESP symbol and (2) its position in a row of five blank spaces. The subjects were thoroughly familiar with ESP symbols as targets but not with position as a target. They got negative scores on the former and positive ones on the latter, with the difference being statistically significant. On the basis of this result, it was predicted that when one of these aspects was novel and the other was familiar, the novel aspect would give the highest score. A further experiment was therefore designed with these two target aspects. The subjects were naïve with regard to both kinds of test material. Half of them were given ten runs apiece with ESP symbols in order to condition them to this target aspect. (Their results on these runs fell below chance to a statistically significant degree.) The other half were given ten runs with just position in the row of five spaces as the target in order to condition them to this feature. Then each subject did two runs in which symbol and position combined were the target. The results did not follow the prediction. Scoring on the dual-aspect targets for both groups was positive on position and negative on symbol. Several tentative explanations for the unexpected nature of the results are offered.

Mr. Stanley J. Hallett, the author, carried out the experiments while he was a member of the Parapsychology Laboratory at Duke University. He is at present a student at Boston University School of Theology.—Ed.

THE RESEARCH to be reported grew out of findings from an analysis by Esther Bond Foster (3) of ESP test data involving multiple-aspect targets. The purpose of Foster's analysis was to determine whether the subjects in the tests had been getting an ESP impression of the target card as a whole or whether they were getting impressions of the separate aspects and were making their complete hits when the ESP hits on the separate aspects happened to coincide. If, for instance, the subject correctly named the Jack of Clubs in an ESP test based on playing cards, did he identify the card as a complete unit or did he get one impression of it as "a Jack" and another as "a Club"? If the latter was what happened, he would, of course, make a complete hit because the two correct identifications of the separate aspects happened to coincide. Foster's study of all the data

available indicated that with multiple-aspect targets the subjects generally obtained an ESP impression of the target as a whole.

These findings led me to plan a new experiment with multiple-aspect targets to see whether the complete hits would be confirmed as a more reliable indication of ESP performance than the single aspects in such tests. I decided to use as the two target aspects (1) the ESP symbol, and (2) its position in a row of five blank spaces, an arrangement suggested by Dr. J. G. Pratt.

Preliminary Experiment

The experiment was conducted in the summer of 1951 with 13 members of the staff of the Parapsychology Laboratory as subjects. Each subject attempted to duplicate the arrangement of ESP symbols on a special target sheet, which was marked off into two columns of one-third inch squares, each column containing a row of five squares horizontally and 25 rows vertically. Thus there was a total of 125 squares in each column and a total of 250 squares on the page. An ESP symbol was placed in one of the five squares in each of the 50 rows. The symbol to be used and the position in the row were determined by a table of random numbers. A different target sheet was made out for each subject, and each was enclosed, without folding, in a Manila envelope. When testing a subject, I, as the experimenter, gave him this envelope with the target sheet face up inside. On the front of the envelope was placed a blank record sheet like that used for the target symbols inside. The subject was instructed to fill out the blank sheet to match the target sheet in the envelope. Each subject filled out one page (two columns) in the presence of the experimenter, who observed that the subject did not open the envelope containing the target sheet.

The general results for the 13 subjects are shown in Table 1. Obviously the multiple-aspect or complete hit results are not outstanding. In fact, there is no deviation at all from the chance level of scoring. The interesting results lie in the fact that the symbol aspect gives a negative deviation and the position aspect gives a positive deviation.

Foster pointed out that it is "theoretically possible that a subject might have different attitudes toward the different aspects of a multiple-aspect target and that these attitudes would result in posi-

Table 1

GENERAL RESULTS SHOWING SCORES ON SYMBOLS, POSITION, AND COMPLETE TARGET IN PRELIMINARY EXPERIMENT

Target	Runs	Probability of Hit per Trial	Expected Number of Hits	Observed Number of Hits	Dev.	CR	P	CR_d
Symbol.....	26	1/5	130	104	−26	2.55	.0108	4.02
Position.....	26	1/5	130	162	+32	3.14	.002	P < .00005
Total.....	52	1/5	260	266	+ 6	.42	.67	
Symbol plus position.... (Complete impression)	26	1/25	26	26	0	0.0	1.00	

tive scoring for one aspect and negative scoring for the other." The difference in scoring rate in the present experiment suggests an effect of novelty upon ESP performance. All the subjects taking the test were accustomed to tests using the ESP symbols, but not to the use of position as the target. Thus the unique part of the test, the position aspect, might have produced a higher rate of scoring because of the favorable effect of novelty.

Pratt and Woodruff (**4**), Thouless (**6**), Cadoret (**2**), and others have stressed the importance of novelty in producing high ESP scoring. Furthermore, past ESP researches involving a comparison of contrasting conditions in the same test (**1, 5, 6**) have indicated that the scores on the favored (in this case the novel) task tended to be positive, while those on the other task tended to fall below the level expected on a chance basis. I decided therefore to design another experiment to test the hypothesis that when two types of targets are used concurrently, the more novel of the two will give the higher scores.

MAIN EXPERIMENT

It seemed wise to use naïve subjects in the experiment rather than to retest the members of the Laboratory, and to control the difference in novelty of the two aspects by the experimental procedure. For this purpose I planned to familiarize or condition the subject on one aspect of the the test (symbols only or position only),

then give him a test which included both the "old" and a new aspect (symbols *plus* position). It was predicted on the basis of the preliminary experiment that in the second part of the test the rate of scoring on the aspect which was novel to the subject would be higher than that on the aspect with which the subject had become somewhat familiar.

There were 40 subjects who participated in this experiment, the majority of whom were visitors to the Laboratory. They were divided into two groups: The A subjects were familiarized with the ESP symbols by writing a page of 250 calls (10 runs) of ordinary DT tests of clairvoyance (these were conditioning trials) previous to doing two dual-aspect runs. The B subjects were familiarized with the position aspect by marking 10 runs of *position* only before doing the two dual-aspect runs. Whether a subject was placed in the A or B group depended upon when he appeared to take the test, the subjects being arbitrarily assigned according to an AABBAABB (etc.) pattern.

The conditioning of the A subjects was accomplished in this manner: Ten target runs were recorded on standard ESP record sheets by an assistant. The selection of these target symbols was based on a table of random numbers, as in the preliminary experiment, and the target sheets were placed in opaque envelopes, in the same manner as in the preliminary experiment, by the assistant, who prepared it from a table of random numbers. Thus I, the experimenter who administered the test, did not know what the target order was and therefore could not give sensory cues to the subject. As in the preliminary experiment a blank record sheet like the target sheet inside was placed on top of the envelope containing the target sheet and it remained there while the subject wrote his calls. I was always present as the subjects took the test and I was careful to observe that they did not open the envelopes containing the target sheets.

For the conditioning of the B subjects, the target page was like the one described in the preliminary experiment, with five squares across and 25 down in each column. The subject was asked to place a check mark (instead of a symbol) in one of the five boxes in each of the fifty rows (two columns) on a page in such a way as to match the target sheet inside the envelope. There were five pages,

each in a separate envelope, for each subject, making a total of 250 trials, or 10 runs. The target squares were selected by the use of a table of random numbers, and the sheets were filled out as above with the same precautions to avoid sensory cues.

The dual-aspect procedure for all subjects was identical with that described in the preliminary experiment and was administered immediately after the subject had completed the conditioning trials.

The results of this part of the experiment are shown in Table 2. The only significant feature of the results is the negative score made by the A subjects when they were taking their conditioning runs with ESP symbols. These 200 runs gave a deviation of −107,

Table 2

RESULTS ON CONDITIONING RUNS AND ON SYMBOL, POSITION, AND COMPLETE IMPRESSION FOR DUAL-ASPECT TARGETS

Target	Runs	Probability of Hit per Trial	Expected Number of Hits	Observed Number of Hits	Dev.	CR	P
A-GROUP							
Symbols (Conditioning runs)	200	1/5	1000	893	−107	3.78	.0002
Symbols (Dual-aspect runs)	40	1/5	200	196	− 4		
Position (Dual-aspect runs)	40	1/5	200	221	+ 21		
Complete impression: Symbol-plus-position (Dual-aspect runs)	40	1/25	40	49	+ 9		
B-GROUP							
Position (Conditioning runs)	200	1/5	1000	1021	+ 21		
Position (Dual-aspect runs)	40	1/5	200	206	+ 6		
Symbols (Dual-aspect runs)	40	1/5	200	186	− 14		
Complete impression: Symbol-plus-position (Dual-aspect runs)	40	1/25	40	37	− 3		

which has a CR of 3.78 (P = .0002). All the other divisions of the table for which no CR's are listed are quite insignificant.

The results do not follow the prediction that the A subjects would score significantly higher on the position aspect of the dual-aspect runs, and that the B subjects would score significantly higher on the symbol aspect in the dual-aspect runs. As is shown by the table, both groups of subjects scored positively on the position aspect of the dual-aspect runs, and negatively on the symbol aspect, although there is not a significant difference between the total hits on position and symbol in the dual-aspect runs.

One of the B subjects had a strikingly high score on the 10 runs of conditioning trials. The deviation of this subject was a +25, giving a CR of 3.96. The probability of having one subject in a group of 40 with a deviation this large or larger is .003. This selected case is not, of course, in itself conclusive evidence of ESP in the experiment, but it takes on added interest against the background of the other significant findings.

Discussion

The most important result of the experiment appears in the 200 runs of conditioning trials for the A subjects. The striking negative deviation in this part of the experiment might be due to several factors. For instance, the subject was brought into the testing room and with a minimum of introductory conversation, was put to work on the test. He knew that the target symbols, which were written in unfamiliar shorthand by the assistant who prepared the target sheets, were small and closely packed together on the target sheet. The laboriousness and difficulty of the task may have motivated him unconsciously to score negatively.

Another, and perhaps less probable, hypothesis might be drawn from the attitude of the experimenter. After the first three or four subjects, I noticed that a negative deviation had begun to accumulate in the conditioning trials. The first four A subjects gave a deviation of −37 for 40 conditioning runs. I was strongly motivated in the direction of having this deviation increase, and there is the possibility that such motivation might have been communicated to the subject by subtle social factors or even by means of extrasensory

perception. However, this would call for an explanation of the negative scoring of the first two or three subjects as being due to chance fluctuations or other factors, which is unlikely for a deviation as large as this: −37 in 40 runs.

It should be noted here that the experimenter would have had no opportunity to obtain such a deviation through motivated recorded error, etc. The target sheets had been prepared by an assistant; they had been placed in the envelopes unseen by the experimenter; the subjects had placed their own calls on the call record sheets; and the calls were checked and rechecked against the targets by persons who marked the hits entirely independently.

In considering the possible reasons for the scoring in opposite directions on the two aspects, it should be noted that there is a considerable difference in the type of mental process that leads to the choices on the different aspects. To select a symbol, the subject must bring the symbol consciously into mind. However, in selecting the position or square in which the check mark is to be placed, it would seem that the choice might be on a more unconscious level, or perhaps more automatic, as all that is required is a simple movement of the hand or arm. Moreover, since the subject's mind was already occupied with the choice of symbol, there is all the more reason to suppose that his choice of position was an unconscious one. Thus it would seem reasonable to speculate that the psi factor might express itself differently in the two aspects. This may account for the fact that the experiment did not confirm Foster's findings regarding complete hits on multiple-aspect targets. In the data she analyzed, the two aspects were more alike from a psychological standpoint, and less artificially linked than in the present case. My experiment, then, may not have provided data that were sufficiently comparable to those she analyzed.

It will be interesting to see whether the use of symbol and position as dual targets in ESP tests will continue to show opposite scoring trends for the two aspects in future experimentation.

References

1. Bevan, J. M. ESP tests in light and darkness. *J. Parapsychol.*, 1947, **11**, 76-89.

2. CADORET, R. J. Effect of novelty in test conditions on ESP performance. *J. Parapsychol.*, 1952, **16**, 192-203.

3. FOSTER, E. B. Multiple aspect targets in tests of ESP. *J. Parapsychol.*, 1952, **16**, 11-22.

4. PRATT, J. G., AND WOODRUFF, J. L. Size of stimulus symbols in extrasensory perception. *J. Parapsychol.*, 1939, **3**, 121-58.

5. STUART, C. E., HUMPHREY, B. M., SMITH, B. M., AND MCMAHAN, E. A. Personality measurements and ESP tests with cards and drawings. *J. Parapsychol.*, 1947, **11**, 118-46.

6. THOULESS, R. H. A comparative study of performance in three psi tasks. *J. Parapsychol.*, 1949, **13**, 263-73.

745 Commonwealth Avenue
Boston 15, Massachusetts

EFFECT OF THE RECEIVER'S ATTITUDE TOWARD THE SENDER IN ESP TESTS

By GEORGE W. CASPER

ABSTRACT: This experiment was an attempt to test the hypothesis that the closer the emotional relationship between sender and receiver in an ESP experiment, the higher the scores. Ten male and ten female college students rated each other on a scale which varied from the one most liked to the one least liked. The experimenter paired off each receiver with one sender of the opposite sex whom he liked most and one whom he liked least. At each session each pair of subjects was given two GESP runs and two BT runs. When all available data were analyzed the results as a whole were found to be statistically significant. Most of this significance was due to the GESP runs, in which it was found that the highest scores were associated with the least liked senders and the lowest scores with the best liked senders. This difference also was statistically significant.

The subjects were rated on the Bernreuter introversion-extraversion scale. These results tended to confirm the conclusion of other experimenters that extraverts score positively and introverts score negatively in ESP tests.—Ed.

INTRODUCTION

THE SUGGESTION has often been made that the degree of emotional attachment of sender and receiver might have an effect upon the scoring level in ESP tests. Stuart, in his ESP experiments with drawings (3) found that couples who had a close relationship (twins and engaged or married couples) made higher scores than did couples who were unrelated. In other cases, such as that of Mrs. Stewart, Dr. Soal's subject, the results were best with agents whom the subject seemed to prefer.

The work herein reported is an attempt to test the hypothesis that the closer the emotional relationship between sender and receiver, the more likely it is that evidence of ESP will be obtained in the experiment. This experiment was done in the psychology laboratory of Heidelberg College, Tiffin, Ohio, in the spring of 1951. The writer was a research assistant at the Parapsychology Laboratory at Duke University at the time, but he returned to Heidelberg to do this experiment because the fine cooperation which the students and faculty had shown in a previous experiment (1) made it an ideal location for further testing.

There were twenty subjects in the experiment, ten men and ten women. They were all college students, mostly upperclassmen, and were personally acquainted with each other. Students were selected to be subjects who were campus leaders or who were generally quite well liked on campus in the opinion of the experimenter and the two faculty members who helped him make the selections.

Procedure

After all the subjects had been selected, each one was given a form containing the names of all twenty who were taking part in the experiment. It was divided into a list of men and a list of women. Each subject was asked to indicate in each list the individual he liked best, the one he liked next best, and so on down to the person he liked least, disregarding his own name as it occurred in the appropriate list. From these forms the experimenter then made a list of couples who were to be tested together, one as the sender and the other as the receiver. Each subject was to be the receiver once with the person of the opposite sex he or she liked most as the sender, and each was to be the receiver once with the person of the opposite sex he liked least as the sender. There were thus forty sessions planned, each one to consist of two runs of GESP and two runs of BT using standard ESP cards.

During the experimental session the experimenter remained with the sender, and the receiver sat alone in a room around the corner of the hall, two rooms distant from the sender. A signal system between rooms was used so that the sender and receiver could keep in time with each other. This signaling system consisted of a small red light and a two-way switch in each room. The sender turned on both his light and the receiver's light with his switch when a card was in position, and the receiver turned off both lights with his switch after recording his call. This signal system was the only communication between sender and receiver during a run.

The sessions were conducted in the following way: After the experimenter had given the sender and receiver preliminary instructions on handling cards and recording calls, the receiver went to his room. The experimenter then shuffled the deck and gave it to the sender, who shuffled it again. The sender then picked up the first card and signalled the receiver to make his first call by turning on

the lights. The receiver, in his room, recorded his call and turned out both lights. Then the sender put the next card in place and switched the lights on again, and the test was continued in this way until a run of 25 cards was completed. Another deck was then shuffled as before, and the test was resumed for another run. Each subject did two runs with the BT procedure, followed by two runs with the GESP procedure, the order being alternated from session to session.

The experimenter, who was seated at a table across from the sender, recorded the cards in a record book with numbered pages and duplicate carbon sheets. In the GESP runs he recorded each card as it was turned over; the BT runs were recorded at the end of each run. The receiver recorded his calls in a similar record book. After each two runs, the receiver returned to the room in which the sender and experimenter were located, so that a check-up of hits could be made. There was no opportunity for cheating on the part of either of the participants at the time of the check-up because the duplicate sheets were removed from the record books by the experimenter and put away before any comparisons were made.

As originally planned there were to be 20 sessions in which the receiver had as a sender the person he liked most in the list of subjects of the opposite sex and 20 sessions in which he had as a sender the person in the list he liked least. However, after about two months had gone by, the experiment was only half completed, so the experimenter called in all the subjects whose work was not done and told them that the experiment had to be finished by a certain date, approximately 2½ weeks away, in order that the experimenter could return to Duke. When the testing ceased on that date, 19 of the 20 sessions with the sender liked most had been completed,[1] but only eight of the 20 sessions with the sender liked least were done. Many of the appointments for the latter sessions had been made, but very often either the sender or the receiver failed to show up for the test. Apparently it was much easier to remember an ESP test with an individual one liked better. All of the results obtained —that is, the data for 27 sessions, or 108 runs—are reported in this paper.

[1] Two of these sessions were done with the same receiver, but with different senders, both of whom the receiver had rated "best liked." These data are included in all the analyses.

Approximately two weeks after the testing was stopped, each subject was asked to fill out the Bernreuter Personality Inventory. These tests were scored to obtain percentile ratings on only the B3-I scale (introversion-extraversion).

Results

The purpose of the experiment was to determine whether or not the differences in attitude of the receiver (in terms of most liked and least liked) would cause him to score at a differential rate with the various senders. The results are therefore broken down into four groups as shown in Table 1. When tested with senders they liked most, the receivers scored below chance on GESP runs (38

Table 1

Total ESP Results

	GESP				BT			
	Runs	Dev.	Av. Run Score	CR	Runs	Dev.	Av. Run Score	CR
Sessions with sender liked most.....	38	−25	4.34	2.03	38	+18	5.47	1.46
Sessions with sender liked least.....	16	+20	6.25	2.50	16	− 5	4.69	.63
Difference in average.....			1.91				.78	
			CR_d=3.19 (P=.0014)				CR_d=1.30 (P=.2)	

Total evaluated by squaring and summing CR's:
χ^2=12.90 (4 d.f.) P=.013

runs with a deviation of −25) and above chance on BT runs (38 runs with a deviation of +18). When tested with the senders they liked least, the receivers scored above chance on GESP runs (16 runs with a deviation of +20), and below chance on BT runs (16 runs with a deviation of −5).

None of the four divisions is independently significant, but when the critical ratios (see Table 1) are squared and summed, a chi square of 12.90 (4 d.f.) is found, with P = .013. The significance is concentrated in the GESP runs, where there is a significant dif-

ference between the results with the two groups of agents (CR_d = 3.19; P = .0014). On the other hand, the difference for the BT groups is not significant.

It is interesting to compare the two ESP procedures in a manner that takes into account the fact that the scoring trends for the two types of tests was opposite for the two groups of agents. To evaluate the distribution as a whole the chi-square test was applied to a fourfold table showing the hits made under each ESP condition with the most liked and the least liked senders. This evaluation based on hits alone was legitimate in this instance since there was the same number of trials with both GESP and BT for each class of agent, and since disregarding the misses only leads to a result that is slightly minimized and is therefore on the safe side. By this method there is a chi-square value of 7.95 (1 d.f.) with an associated probability of .005. The results as a whole, then, are clearly significant at the one percent level.

Another analysis was done in this experiment. The 20 subjects were divided into two groups: the introverts (those scoring above the fiftieth percentile on the B3-I scale of the Bernreuter) and the extraverts (those scoring below the fiftieth percentile). The introverts scored below chance (28 runs with a deviation of −18) and the extraverts scored above chance (80 runs with a deviation of +26). This difference is suggestive of an extrachance distribution, the CR_d being 2.21.

The question of the relation between ESP scoring and introversion-extraversion may also be asked in terms of how the 20 subjects were divided. The answer is shown in the following fourfold table:

	ESP Deviation		
	+	0 or −	Total
Extraverts	8	6	14
Introverts	0	6	6
Total	8	12	20

The smallest numbers in the cells are below the limit required for the use of the chi-square method. Therefore the distribution has been evaluated by the method for obtaining the exact probability of finding this particular distribution with these marginal totals. (In this case there is no more extreme distribution.) The value of P by this method is found to be .024, which approaches significance in spite of the small number of subjects involved, and the findings of this experiment tend to confirm those of Humphrey (2) on the relation of extraversion and positive scoring.

Discussion

There is little that can be added to the results themselves at this time in the way of an interpretation. The experiment as a whole gives evidence of ESP. Not only were the deviations in the four groupings of the data significant when evaluated jointly, but the distribution of hits in relation to the two test procedures and the most liked and least liked senders was also significant.

The experiment was designed with the expectation that subjects would score more positively with the senders they liked most than with those they liked least. It was supposed that the person who served as sender during the GESP runs would be only an inactive or neutral bystander during the BT runs and that these BT runs would therefore show no difference based on the receiver's attitude toward him. It is clear from the results that the findings did not follow this expected pattern in all respects. The difference shown by the BT runs was not large enough to justify any other assumption than that they were neutral; but the GESP runs were positive with the least liked senders and negative with the most liked senders. At the same time, the fact that most of the significance of the experiment as a whole was contributed by the GESP trials indicates that the sender's looking at the cards was a factor in the receiver's scoring.

Perhaps the most interesting finding at this stage of our knowledge of ESP is the evidence tending to confirm the finding that extraverts score positively in ESP tests while introverts score below the level of chance expectation.

REFERENCES

1. CASPER, G. W. A further study of the relation of attitude to success in ESP scoring. *J. Parapsychol.*, 1951, **15**, 139-45.
2. HUMPHREY, B. M. Introversion-extraversion ratings in relation to scores in ESP tests. *J. Parapsychol.*, 1951, **15**, 252-61.
3. STUART, C. E. GESP experiments with the free response method. *J. Parapsychol.*, 1946, **10**, 21-35.

99 Brattle Street
Cambridge, Massachusetts

NOTES AND COMMENTS

EXPERIMENTS IN ESP AND PK WITH ABORIGINAL SUBJECTS

By Ronald Rose

In June, 1951, my wife and I reported the results of a series of GESP and PK tests carried out with detribalized aborigines at Woodenbong, New South Wales (see *Journal of Parapsychology*, June, 1951). The GESP portion yielded highly significant results, largely because of one high-scoring subject. A second short series conducted a little later with a similar group of natives at Tabulam, New South Wales, produced results in both GESP and PK that were quite insignificant.

From December, 1950, to February, 1951, my wife and I visited groups of natives in Central Australia, principally at Hermannsburg (approximately 100 miles from Alice Springs) and in the Areyonga Valley.[1] The natives at Hermannsburg were, in the main, detribalized Aranda (Arunta) people, and those at Areyonga were almost fully tribal members of the Pitjendadjara tribe.

With these natives we conducted a third series of psi tests using experimental methods that were the same as those reported earlier and test conditions not greatly different. In the GESP tests a total of 25 native subjects participated. The 171 runs which they carried out gave a deviation of 52 above mean chance expectation. This yields a CR of 1.99, which is not statistically significant. An interesting difference between the scoring of the tribal and the missionary groups was found. The missionary natives completed 123 runs with a deviation of only +11 (CR = .49), whereas the tribal natives, doing 48 runs, obtained a deviation of +41, which yields a CR of 2.96. The difference between the scoring of the two groups is not quite significant (CR = 2.24).

A total of 1,128 PK runs with 12 dice per throw (using dice supplied by the Parapsychology Laboratory at Duke University)

[1] We wish to record our appreciation of the cooperation of the Department of Native Affairs, Darwin, and the Lutheran Mission in making these tests possible.

produced an insignificant positive deviation of only +7. All faces of the dice were used as target an equal number of times for each subject. Only one subject produced results that at any time appeared encouraging. His first session of 24 runs showed a deviation of +26 (CR = 2.91), but his final deviation in 144 runs was only +30 (CR = 1.37).

Aboriginal "clever men" (native doctors, magical practitioners) showed no special ability in the experimental situation.

A summary, at this stage, of all our results with aborigines might be of some interest and might serve to place on record the series that do not in themselves justify detailed reports. In all, my wife and I have carried out GESP tests totaling 13,150 calls and PK tests totaling 76,608 individual die-readings with aboriginal subjects. The accompanying table gives these results. Although the results at

COMPLETE RESULTS OF THREE EXPERIMENTS WITH AUSTRALIAN ABORIGINES

GESP TESTS

Location	Runs	Hits	Dev.	Av. Score	CR
Woodenbong	296	1706	+226	5.76	6.57
Tabulam	59	319	+ 24	5.41	1.56
Central Australia	171	907	+ 52	5.30	1.99
Total	526	2932	+302	5.57	6.60

PK TESTS

Location	Runs	Hits	Dev.	Av. Score	CR
Woodenbong	1896	7712	+128	4.07	1.61
Tabulam	168	645	− 27	3.84	1.14
Central Australia	1128	4519	+ 7	4.01	.11
Total	3192	12876	+108	4.03	1.1

Tabulam and Central Australia were without statistical significance in themselves, the trend of scoring continued and the critical ratio established in the first series was maintained.

Camp Street
Toowong, Brisbane, S.W. 1
Queensland, Australia

GLOSSARY

In order to avoid constant redefining of commonly recurring terms in papers appearing in this JOURNAL, the following definitions are submitted for convenient reference.

A simple description of the basic experimental methods and of the statistical procedures to be followed in evaluating ESP (both card tests and drawings tests) and PK may be found in the *Handbook of Tests in Parapsychology* by Betty M. Humphrey (published by the Parapsychology Laboratory, Duke University, Durham, N. C.).

AGENT: In tests for telepathy, the person whose mental states are to be apprehended by the percipient. In GESP tests, the person who looks at the stimulus object.

AVERAGE SCORE: Average number of hits per run.

BT ("Broken Technique"): The clairvoyance technique in which each card is laid aside by the experimenter as it is called by the subject. The check-up is made at the end of the run.

CALL: The ESP symbol selected by the subject in trying to identify a target.

CHANCE: The complex of undefined causal factors irrelevant to the purpose at hand.

CHANCE EXPECTATION = MEAN CHANCE EXPECTATION: The most likely score if only chance obtains.

CHANCE AVERAGE: Mean chance expectation expressed as an expected score, generally in terms of average per run.

CHI SQUARE: A sum of quantities each of which is a deviation squared divided by an expected value. Also a sum of the squares of CR's.

CLAIRVOYANCE: Extrasensory perception of objective events as distinguished from telepathic perception of the mental state of another person.

CR (CRITICAL RATIO): A measure to determine whether or not the observed deviation is significantly greater than the expected random fluctuation about the average. The CR is obtained by dividing the observed deviation by the standard deviation. (The probability of a given CR may be obtained by consulting tables of the probability integral, such as Pearson's.)

CR OF THE DIFFERENCE: The observed difference between the average scores of two samples of data divided by the standard deviation of the difference.

DECK: Twenty-five ESP cards.

CLOSED DECK: An ESP deck composed of five each of the five symbols.

OPEN DECK: An ESP deck made up of the ESP symbols selected in random order.

DEVIATION: The amount an observed number of hits or an average score varies from the mean chance expectation or chance average. A deviation may be total (for a series of runs) or average (per run).

DIE THROW: The elementary unit of data in a dice-throwing test represented by the throwing and reading of a single die, no matter how many dice are thrown at a time.

DISPLACEMENT: ESP responses to targets other than those for which the calls were intended.

BACKWARD DISPLACEMENT: ESP responses to targets preceding the assigned targets (the ones for which they were intended). Displacement to the targets

one, two, three, etc. places preceding the assigned target are designated as (—1), (—2), (—3), etc.

FORWARD DISPLACEMENT: ESP responses to targets coming later than the assigned targets. Displacement to the targets one, two, three, etc. places after the assigned target are designated as (+1), (+2), (+3), etc.

DT (Down Through): The clairvoyance technique in which the cards are called down through the deck before any are removed or checked.

EMPIRICAL CONTROL: An experiment which wholly or partially follows the main experiment with the exception that the conditions are designed to exclude the possibilty of the hypothesis being tested.

ESP (EXTRASENSORY PERCEPTION): Response to an external event not presented to any known sense.

ESP CARDS: Cards, each bearing one of the following five symbols: star, circle, three parallel wavy lines (called "waves"), square, plus.

ESP SYMBOLS: See plate opposite page 1, this JOURNAL, Vol. 1, March, 1937.

EXPECTATION; see CHANCE.

EXTRACHANCE: Not due to chance alone.

FREE MATERIAL: The stimulus objects of experiments in which an unlimited or unspecified range of stimulus objects is employed (as contrasted with methods such as card-calling in which the subject knows that the stimulus object is one of a known range).

FREE RESPONSE METHOD: An ESP test method in which a relatively unlimited range of stimulus objects is used and in which the subject is instructed to express freely as his response any range of thoughts and ideas.

GESP (GENERAL EXTRASENSORY PERCEPTION): A technique designed to test the occurrence of extrasensory perception, permitting either telepathy or clairvoyance or both to operate.

HIGH-DICE TEST: A PK technique in which the aim of the subject is to try to influence a pair of dice to fall with the two upper faces totaling eight or more.

LOW-DICE TEST: A PK technique in which the aim of the subject is to try to influence a pair of dice to fall with the two upper faces totaling six or less.

MEAN CHANCE EXPECTATION; see CHANCE.

P (PROBABILITY): A mathematical estimate of the expected relative frequency of a given event if chance alone were operative.

PARAPSYCHOLOGY: A division of psychology dealing with those psychical effects which appear not to fall within the scope of what is at present recognized law.

PERCIPIENT: The person who makes the calls in an ESP test.

PK (PSYCHOKINESIS): The direct influence exerted on a physical system by a subject without any known intermediate physical energy or instrumentation.

PRECOGNITION: Cognition of a future event which could not be known through rational inference.

PREFERENTIAL MATCHING: A method of scoring free responses. A judge ranks the stimulus objects (usually in sets of four) with respect to their similarity to, or association with, each response; and/or he ranks the responses with respect to their similarity to, or association with, each stimulus object.

PSI: A general term to identify personal factors or processes in nature which transcend accepted laws. It approximates the popular use of the word "psychic" and the technical one, "parapsychical."

PSI PHENOMENA: Occurrences which result from the operation of psi. They include the phenomena of both ESP (including precognition) and PK.

QD (QUARTER DISTRIBUTION): The distribution of hits in the record page (or in a logical subdivision thereof, such as the set or the half-set) as found in the four equal quarters formed by dividing the selected unit horizontally and vertically.

RUN: A group of trials, usually the successive calling of a deck of 25 ESP cards or symbols. In PK tests, 24 single die throws regardless of the number of dice thrown at the same time.

SCORE: The number of hits made in one run.

TOTAL SCORE: Total of scores made in a given number of runs.

AVERAGE SCORE: Total score divided by number of runs.

SD (STANDARD DEVIATION): Usually the theoretical root mean square of the deviations. It is obtained from the formula $\sqrt{npq}$, in which n is the number of single trials, p the probability of success per trial, and q the probability of failure. (For ESP cards, SD $= 2\sqrt{\text{no. of runs.}}$) Sometimes the SD used is that estimated from the observed variability in the scores.

SD OF THE DIFFERENCE: For both ESP cards and PK tests using dice, the SD of the difference is equal to $\sigma_s \sqrt{1/R_1 + 1/R_2}$ where σ_s is the SD of a single run and R_1 and R_2 are the number of the runs in the respective samples compared. This gives the SD of the difference for run score averages.

SERIES: Several runs or experimental sessions that are grouped in accordance with a stated principle.

SESSION: A unit of an ESP or PK experiment comprising all the trials of one test occasion.

SET: A subdivision of the record page serving as a scoring unit for a consecutive group of trials, usually for the same target.

SEVENS TEST: A PK technique in which the aim of the subject is to try to influence a pair of dice to fall with the two upper faces totaling seven.

SIGNIFICANCE: A numerical result is significant when it equals or surpasses some criterion of degree of chance improbability. The criterion commonly used in this JOURNAL is a probability value of .02 or less, or a deviation in either direction such that the CR is 2.33 or greater.

SINGLES TEST: A PK technique in which the aim of the subject is to try to influence dice to fall with a specified face up.

STIMULUS OBJECT: The ESP card or drawing or other object, some identifying characteristic of which is to be apprehended by the percipient.

SUBJECT: The person who is experimented upon. In ESP tests, most commonly the percipient (though also the agent in GESP and telepathy). In PK tests, any individual whose task it is to influence the objects thrown.

TARGET: In ESP tests, the stimulus object; or in telepathy, the mental state of the agent. In PK tests, the faces of the die (or combination of faces) which the subject attempts to bring up in the act of throwing.

TARGET CARD: The card which the percipient is attempting to perceive (i.e., to identify or otherwise indicate a knowledge of).

TARGET DECK: The deck of cards the order of which the subject is attempting to identify.

TARGET FACE: The face on the die which the subject tries to turn up as a consequence of direct mental action.

TELEPATHY: Extrasensory perception of the mental activities of another person. It does not include the clairvoyant perception of objective events.

TRIAL: In ESP tests, a single attempt to identify a stimulus object. In PK tests, a single throw of the dice or other objects thrown.

SHOULD SCIENTISTS EXAMINE EACH OTHER'S MOTIVES?

The disclosures of modern psychiatry and psychology regarding human motivation have placed the scientist in a vulnerable position. He must now admit—however reluctantly—that unconscious emotional influences may color his scientific perspective despite all his efforts at objectivity. This ought to make him more chary than ever of departing from the rigorous criteria for evidence and proof that science has established. On the other hand, this new knowledge presents him with a constant temptation to abandon the traditional methods of logical argument and experimental proof in controversial matters for an analysis of the intents and motives of his intellectual adversaries. The presumption underlying this approach to polemics is that if one can question the motives behind an individual's viewpoints, it is not necessary to consider seriously the viewpoint itself. No one will fail to recognize that the genetic fallacy is inherent in this position, and while this has been thoroughly discredited as a logical procedure, it nevertheless has a strong appeal. The lure is perhaps greatest for those engaged in the psychological disciplines since they are especially conversant with the principles of dynamic psychology and particularly sensitive to the frailty of human reason.

* * * *

The revelations of dynamic psychology are a heady wine, and psychologists must beware of losing their sobriety. Science provides the methods for discovering the truth or falsity of an investigator's views. If these methods are incapable of resolving an issue, it is sanguine to expect an examination of individual motives to do so. On the contrary, the history of ideas has shown that the injection of personal issues only leads to dissension and the growth of intensely partisan schools. Vituperation takes the place of objective discussion, and the truth is obscured by the dust of battle.

—Letter to the Editor by Samuel Waldfogel
(Massachusetts General Hospital)
American Psychologist, May, 1952
(Publication of the American Psychological Association)